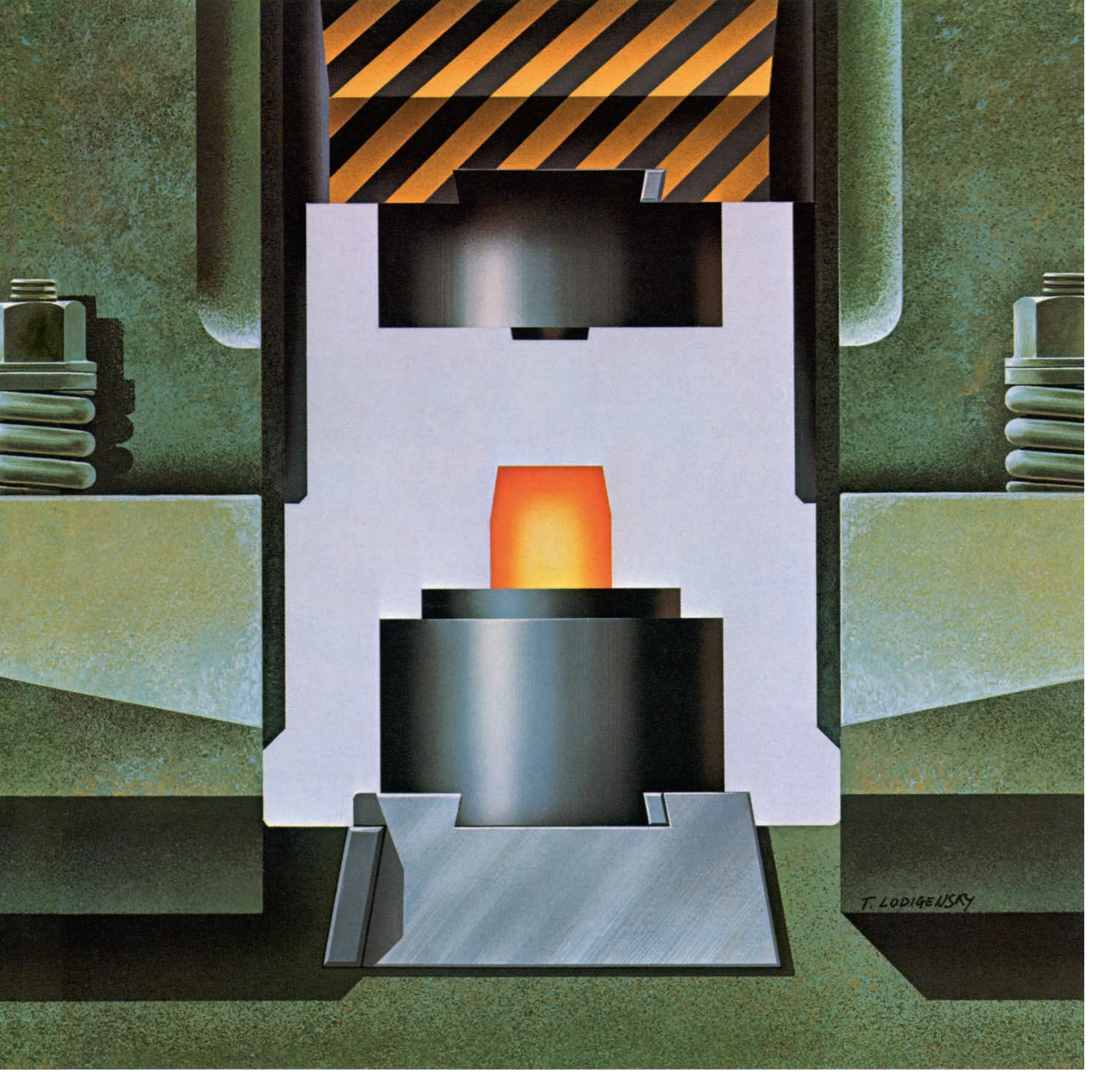
INVESTIGACION Y CIENCIA

Edición en español de SCIENTIFIC AMERICAN



FORJA

Octubre 1981

300 Ptas.

Los espacios en gris
corresponden a publicidad
en la edición impresa

Número 61
Octubre 1981

INVESTIGACION Y
CIENCIA
Edición en español de SCIENTIFIC AMERICAN

EDITA Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (ESPAÑA)

Colaboradores de este número:

Asesoramiento y traducción:

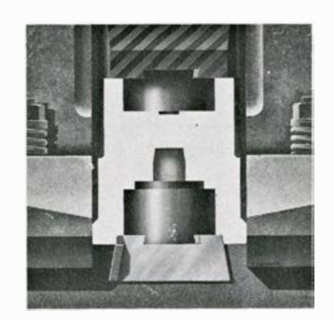
Guillermo Jenaro Garrido: *Armas con guía de precisión*; Manuel Puigcerver: *Las estrellas más recientes de Orión*; Ricardo Sánchez Tamés: *La enfermedad holandesa del olmo*; Santiago Torres Martínez: *El ribosoma*; Pedro Molera Solá: *La forja de los metales*; Manuel Puigcerver: *Los monzones*; Basilio Losada Castro: *El programa del alcohol en Brasil*; Víctor Navarro Brotons: *Sadi Carnot*; Luis Bou: *Juegos matemáticos*; J. Vilardell: *Taller y laboratorio*.

Ciencia y sociedad:

J. P. Garrido Roiz y Mercedes Molina Ibáñez

Libros:

P. Laín Entralgo, Antonio Tovar y Juan Vernet



LA PORTADA

La ilustración de la portada representa de un modo esquemático un martillo de forja (véase "La forja de los metales", por S. L. Semiatin y G. D. Lahoti). El pesado mango de acero de la parte superior, señalado con franjas de seguridad, lleva una matriz y se mueve en sentido descendente. La pieza de trabajo caliente (el metal que va a ser forjado) permanece sobre una matriz fijada a la parte inferior. Los resortes permiten absorber la energía de los golpes producidos por el mango. Tales martillos, algunos de los cuales son impulsados por gravedad y otros con auxilio de energía, vapor o aire a presión, son las máquinas de forja más antiguas, rápidas y versátiles. Un martillo trabaja golpeando la pieza repetidamente. Las prensas mecánicas e hidráulicas hechuran la pieza en una sola operación. En una prensa mecánica, la energía se almacena en una rueda de inercia y el mango se acciona por medio de una manivela o una leva. En una prensa hidráulica, el mango es accionado mediante un fluido a presión elevada.

Suscripciones:
Prensa Científica, S. A.
Calabria, 235-239
Barcelona-29 (España)
Teléfono 322 05 51 ext. 41

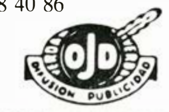
Condiciones de suscripción:
España:
Un año (12 números): 3.300 pesetas
Extranjero:
Un año (12 números): 52 U.S. $
Ejemplar atrasado ordinario:
340 pesetas
Ejemplar atrasado extraordinario:
440 pesetas

Distribución para España
Distribuciones de Enlace, S. A.
Ausias March, 49, Barcelona-10

Distribución para los restantes países:
Editorial Labor, S. A.
Calabria, 235-239 - Barcelona-29

Publicidad:
Madrid:
Gustavo Martínez Ovin
Avda. de Moratalaz, 137, Madrid-30
Tel. 430 84 81
Cataluña:
Miguel Munill
Balmes, 191, 2.°, 2.ª, Barcelona-6
Tels. 218 44 45 y 218 40 86

Controlado por O.J.D.



PROCEDENCIA DE LAS ILUSTRACIONES

Pintura de la portada de Ted Lodigensky

Página	Fuente
10	Departamento de Defensa de los EE.UU.
12-13	Andrew Christie
14-15	Vought Corporation
16-19	Andrew Christie
23	Observatorio del Harvard College
24	Observatorio Nacional Kitt Peak (*izquierda*) Nicholas Z. Scoville, F. Peter Schloerb y Paul F. Goldsmith (*derecha*)
25	Observatorio Lick
26-31	Allen Beechel
35-37	Tom Prentiss
38	Ruud J. Scheffer, Laboratorio Willie Comelin Scholten
39	Dockle M. Elgersma, Laboratorio Willie Comelin Scholten
40	Tom Prentiss
41	Tom Prentiss (*arriba*) Gerald N. Lanier (*abajo*)
42	Tom Prentiss
43	Gerald N. Lanier (*arriba*) Bruce Hemming (*abajo*)
44-45	Gary A. Strobel
47-48	Juan Pedro Garrido Roiz
50-51	Mercedes Molina Ibáñez
54-55	George V. Kelvin (*izquierda*) James A. Lake (*derecha*)
56	James A. Lake (*arriba*) George V. Kelvin (*abajo*)
57	George V. Kelvin
58	James A. Lake y Lawrence Kahan
59-65	George V. Kelvin
66	Michael Clark y James A. Lake
73	A. A. Badawy, C. F. Billhardt y W. J. Weider
74-75	Gabor Kiss
76	W. J. Weider
77	Gabor Kiss (*arriba*), C. R. Thompson (*abajo*)
78	W. W. Sunderland y Gabor Kiss (*arriba*), C. R. Thompson y Gabor Kiss (*abajo*)
79	W. W. Sunderland y Gabor Kiss
80	Mesta Machine Company
84	National Oceanic and Atmospheric Administration
86-94	Alan D. Iselin
97	Miguel Alonso
98-104	José Roberto Moreira y José Goldemberg
106	J. L. Charmet, Collection Particulière, Paris
108	Science and Technology Research Center, New York Public Library, fundaciones Astor, Lenox y Tilden
109-110	Ilil Arbel
111-116	Stuart L. Rabinowitz
127	Jearl Walker
128-130	Michael Goodman

ISSN 0210-136X
Dep. legal: B. 38.999-76
Fotocomposición Tecfa
Pedro IV, 160 - Barcelona-5
Fotocromos reproducidos por GINSA, S.A.
Imprime GRAFESA
Gráfica Elzeviriana, S. A.
Nápoles, 249 - Tel. 207 40 11
Barcelona-13
Printed in Spain - Impreso en España

Los autores

PAUL F. WALKER ("Armas con guía de precisión") es director de publicaciones sobre investigación de control de armamentos de la "Union of Concerned Scientists". En 1969, tras su primer ciclo universitario en el College of the Holy Cross, se alistó en el ejército de los Estados Unidos. Allí se hizo traductor de ruso y se especializó en los servicios de información. En 1973 terminó su licenciatura en relaciones internacionales, cursada en la Facultad de Estudios Internacionales Avanzados de la Universidad Johns Hopkins. Carrera que culminó en 1978 doctorándose en el Instituto de Tecnología de Massachusetts. Disfrutó de una beca de investigación postdoctoral en el Centro de Ciencia y Asuntos Internacionales de la Facultad John F. Kennedy de la Universiad de Harvard. Escritor, docente y consultor en materias relativas a la seguridad nacional, control de armamentos y relaciones ruso-americanas, Walker fue uno de los seis coautores del libro: *The price of defense: A new Strategy for Military Spending*, publicado por Times Book, una de las cuatro divisiones de The New York Times Book Company, Inc.

GARETH WYNN-WILLIAMS ("Las estrellas más recientes de Orión") es profesor agregado de astronomía y director del programa de estudios superiores de astronomía de la Universidad de Hawai en Manoa. Nacido en Londres, obtuvo su licenciatura en ciencias naturales por la Universidad de Cambridge, en 1966. Explicó física durante un año en el Colegio Universitario de Nairobi antes de regresar a Cambridge para trabajar como postgraduado en astronomía bajo la dirección de Martin Ryle. Fue elegido miembro del Trinity College en 1969 y obtuvo su doctorado en 1971. Desde 1971 hasta 1973 realizó trabajo postdoctoral en el Instituto de Tecnología de California, y, de 1973 a 1978, formó parte del grupo de radioastronomía del Cavendish Laboratory, de Cambridge. Wynn-Williams marchó a Hawai hace tres años.

GARY A. STROBEL y GERALD N. LANIER ("La enfermedad holandesa del olmo") trabajan en un mismo empeño, a pesar de la distancia geográfica que separa sus centros de investigación; están especializados en diferentes aspectos de la biología y control de la enfermedad holandesa del olmo. Strobel ocupa la cátedra R. G. Gray de fitopatología de la Universidad estatal de Montana. Se doctoró en fitopatología por la Universidad de California en Davis, en 1963. Se ocupa de la bioquímica de las interacciones entre parásitos vegetales y la planta huésped. Lanier es profesor de entomología en Syracuse, en la Facultad de Ciencias del Medio Ambiente y Silvicultura de la Universidad estatal de Nueva York. Cursó toda su carrera universitaria en Berkeley, coronándola con el doctorado que obtuvo allí en 1967. El tema de su tesis fue la ecología química de los escarabajos de la corteza. Antes de trasladarse a Nueva York, Lanier investigó en el Laboratorio Canadiense de Silvicultura de Calgary, Alberta.

JAMES A. LAKE ("El ribosoma") es profesor de biología molecular en la Universidad de California en Los Angeles. Licenciado en física por la de Colorado en Boulder, en 1963, recibió el doctorado, cuatro años más tarde, en la de Wisconsin en Madison. "Hasta que no acabé mi carrera de física teórica no me decidí a centrar mis esfuerzos en la biología molecular. Pensé que comprender los mecanismos moleculares de la vida me reportaría más satisfacciones que seguir dedicándome a la física. Desde entonces, mi investigación se ha polarizado en la estructura y función del ribosoma." Ha trabajado en el Instituto de Tecnología de Massachusetts, la Facultad de Medicina de Harvard, la Universidad Rockefeller y la Facultad de Medicina de la Universidad de Nueva York. En 1975, Lake obtuvo el primer Premio Burton de la Sociedad Americana de Microscopía Electrónica, por su trabajo sobre la estructura del ribosoma.

S. L. SEMIATIN y D. G. LAHOTI ("La forja de los metales") son expertos en tecnología de la metalistería. Trabajan en el Battelle Memorial Institute de Columbus, Ohio. Semiatin estudió en la Universidad Johns Hopkins y en la Carnegie-Mellon. Su campo de investigación se centra en las propiedades de formación de los metales. Lahoti se educó en la Universidad de Burdewan y en la de California en Berkeley. Se ha especializado en los modelos matemáticos de los procesos de la forja de metales. Desde que entró en Battelle ha trabajado en varios proyectos que habrían de acercar a la industria las técnicas de simulación con ordenadores.

PETER J. WEBSTER ("Los monzones") es jefe de investigación de la división de física atmosférica de la Organización de la Commonwealth para la Investigación Científica e Industrial (CSIRO) de Australia. Pertenece al estamento docente del departamento de matemáticas de la Universidad Monash en Melbourne. Inglés de nacimiento, Webster comenzó trabajando de meteorólogo, en 1961, en la Oficina de Meteorología de la Commonwealth en Australia. Estudió física durante un año en el Real Instituto de Tecnología de Melbourne, antes de matricularse en segundo ciclo en el Instituto de Tecnología de Massachusetts, donde recibió el doctorado en meteorología en 1971. Explicó ciencias atmosféricas en la Universidad de California en Los Angeles y en la Universidad de Washington hasta 1977, regresando entonces a Australia. Especialista en meteorología tropical, ha realizado también investigación en atmósferas planetarias y climas antiguos. Como miembro del Comité del Experimento Monzónico de la Academia Nacional de Ciencias, organizó y dirigió la primera fase de este extenso proyecto en 1978 y 1979.

JOSE R. MOREIRA y JOSE GOLDEMBERG ("El programa del alcohol en Brasil") comparten una misma labor profesoral en el Instituto de Física de la Universidad brasileña de São Paulo. Moreira es ingeniero eléctrico por la Escuela Politécnica de la Universidad de São Paulo, centro superior donde se doctoró en 1971. Pasó largas temporadas de investigación y docencia fuera de su país. Goldemberg dirige el Instituto de Física arriba mencionado. Ha enseñado en las universidades de Stanford, París y Toronto.

S. S. WILSON ("Sadi Carnot") es profesor de ingeniería en la Universidad de Oxford. Allí se graduó en 1944 y se incorporó a su facultad en 1946. Experto en termodinámica, fue responsable del diseño y equipamiento del actual laboratorio de máquinas térmicas de Oxford. Sus investigaciones, según nos informa, conciernen al uso de fluidos pesados en ciclos termodinámicos basados en el vapor, un desarrollo que él considera alternativo al vapor de agua en instalaciones de poca potencia. Además, se ocupa activamente del diseño de máquinas a pedal de aplicación estacionaria y para el transporte, tanto en los países en vías de desarrollo como en los desarrollados. Uno de sus diseños es un chasis "rickshaw" (cochecito japonés de dos ruedas tirado por hombres) mejorado llamado el oxtrike.

Hace…

José M.ª López Piñero

… doscientos años

Félix de Azara marchó a la América meridional, que durante las dos décadas siguientes sería el escenario de su gran obra de naturalista.

Había nacido en 1746 en Barbuñales, localidad aragonesa cercana a Barbastro. Su primera formación discurrió en la Universidad de Huesca, donde cursó estudios que hoy llamaríamos secundarios desde 1757 a 1761. Ingresó en el ejército en 1764 y, un año más tarde, pasó a Barcelona a estudiar matemáticas e ingeniería en la academia militar allí existente, de la que salió en 1767 con el título de «ingeniero delineador de los ejércitos nacionales, plazas y fronteras» y el grado de subteniente. Durante los trece años restantes trabajó como ingeniero militar, participando, entre otras tareas, en la fortificación de la plaza de Figueras, en obras hidráulicas en los ríos Jarama y Henares y en la reconstrucción y reparación de las fortalezas de Mallorca. El año 1775 tomó parte en la expedición contra Argel y, en el desembarco, sufrió una grave herida que estuvo a punto de costarle la vida. Tras permanecer algún tiempo inactivo, se incorporó en 1780 a la guarnición de San Sebastián, con el empleo de teniente coronel de ingenieros.

A comienzos de 1781, Azara recibió órdenes de presentarse ante el embajador español en Lisboa. Este le comunicó que tenía que marchar inmediatamente a América del Sur como miembro de una comisión encargada, junto a otra similar portuguesa, de rectificar los límites entre Brasil y las colonias españolas vecinas, tal como había sido acordado en el tratado de San Ildefonso.

Tras una corta estancia en Río de Janeiro, los comisionados llegaron a Montevideo, donde el virrey español les dio las instrucciones relativas a su trabajo. Azara estuvo en las cercanías de Buenos Aires hasta mediados de 1783, fecha en la que fue enviado a San Pedro, capital del territorio brasileño de Rio Grande do Sul, donde conferenció con las autoridades militares portuguesas. A su regreso fue enviado a Asunción, ciudad que fue su principal residencia durante casi trece años, a lo largo de los cuales realizó numerosos viajes por tierras paraguayas. En 1796, el virrey le encargó inspeccionar la frontera meridional del Río de la Plata y, más tarde, diversas misiones en el río Paraná y en la zona limítrofe con Brasil. A finales de 1801, casi veinte años después de su partida desde Lisboa, regresó a España.

La tarea primordial de Azara durante todo ese tiempo había sido de índole cartográfica: "El principal objeto de mis viajes —afirma en uno de sus libros— era levantar la carta exacta de aquellas regiones, porque ésta era mi profesión y tenía los instrumentos necesarios." No obstante, muy pronto asoció a dicho trabajo la observación de la naturaleza de la zona, hasta entonces muy poco conocida por los científicos europeos. A comienzos de 1790 comenzó a redactar una exposición detenida de sus observaciones, tras haber abandonado la anotación diaria de las mismas, por desordenada y repetitiva. Reformada al menos un par de veces, esta exposición fue publicada más tarde con el título de *Viajes por la América meridional*. Su contenido es muy amplio, ya que incluye "la descripción geográfica, política y civil del Paraguay y del Río de la Plata, la historia del descubrimiento y conquista de estas regiones, detalles numerosos sobre su historia natural y sobre los pueblos salvajes que las habitan, el relato de los medios empleados por los jesuitas para someter y civilizar a los indígenas, etc.". Sin embargo, Azara no se quedó en naturalista indiferenciado, al modo de los grandes estudiosos españoles de la naturaleza americana durante el siglo XVI. Sus principales contribuciones fueron posibles porque se especializó en dos aspectos concretos de la fauna del territorio: "Como estaba yo solo y los objetos que veía eran muchos más de los que podía examinar, me ví precisado a preferir, después de lo dicho, la descripción de los pájaros y cuadrúpedos, quedándome pocos momentos para reflexionar sobre las tierras, piedras, vegetales, pescados, insectos y reptiles."

Sus investigaciones sobre las aves las comenzó al poco de llegar a América, porque en agosto de 1783 ya había terminado el primer volumen de su obra acerca del tema. El material se enriqueció extraordinariamente durante sus años de estancia en Paraguay, hasta reunir en tres volúmenes descripciones extraordinariamente precisas de 448 aves diferentes, en su mayor parte correspondientes a especies o subespecies distintas. Azara estudió, además, con gran cuidado, sus hábitos de alimentación y nidificación, sus hábitats y su distribución geográfica. A partir de 1783, se ocupó de forma parecida de los "cuadrúpedos" existentes en la zona, tanto de los que eran análogos a los del Viejo Mundo, como de los peculiares del lugar.

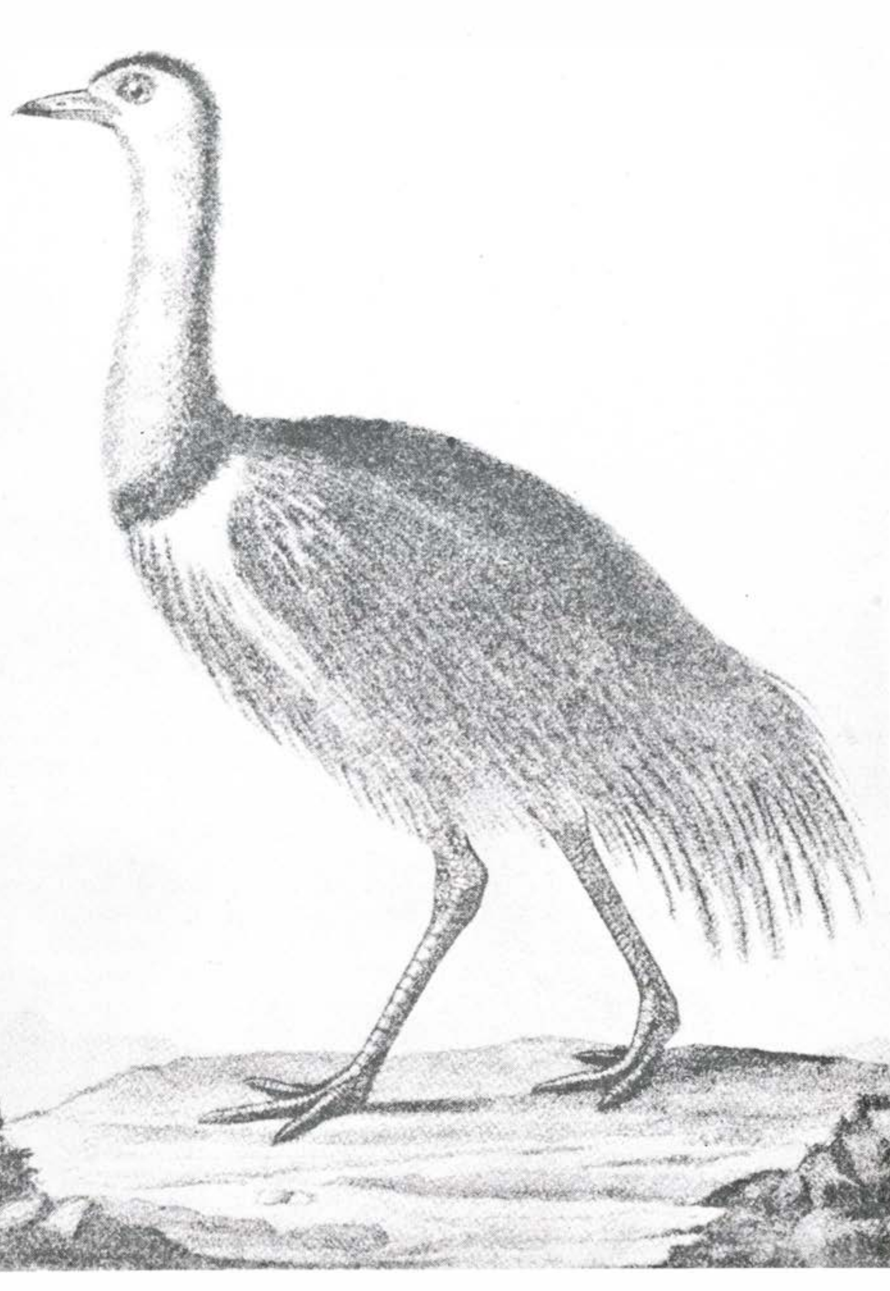
Cuando llegó a Sudamérica, Azara carecía de una preparación específica de naturalista. Desconocía incluso la terminología utilizada por los zoólogos de la época, de forma que al principio ideó una propia, anotando en una lista los significados de los vocablos que empleaba en sus descripciones. En una de sus estancias en Buenos Aires consiguió un ejemplar de la *Histoire naturelle* de Buffon, los doce primeros volúmenes en la traducción castellana de José Clavijo Fajardo y los diecinueve restantes, con sus suplementos, en el original francés. Su admiración por esta gran obra no le impidió darse cuenta de sus limitaciones y errores: "Encontré que buena parte de lo que es histórico se componía de noticias vulgares, falsas o equivocadas; que en lo general no se daba idea exacta de las magnitudes, ni de las proporciones; que se reunían a veces bestias diferentes, embrollándolas; que en ocasiones se multiplicaban las especies…" Creyendo que Buffon vivía aún, Azara pensó al principio escribir una especie de complemento de su obra que proyectaba llevarle personalmente a París.

Azara comenzó a superar su aislamiento científico inicial cuando entró en relación con Antonio de Pineda y Ramírez, uno de los naturalistas que figuraban en la expedición científica dirigida por Alejandro Malaspina. Pineda revisó las observaciones de Azara, corrigió algunos errores y repeticiones, animándole al mismo tiempo a continuar su trabajo.

Antes de volver a España, Azara en-

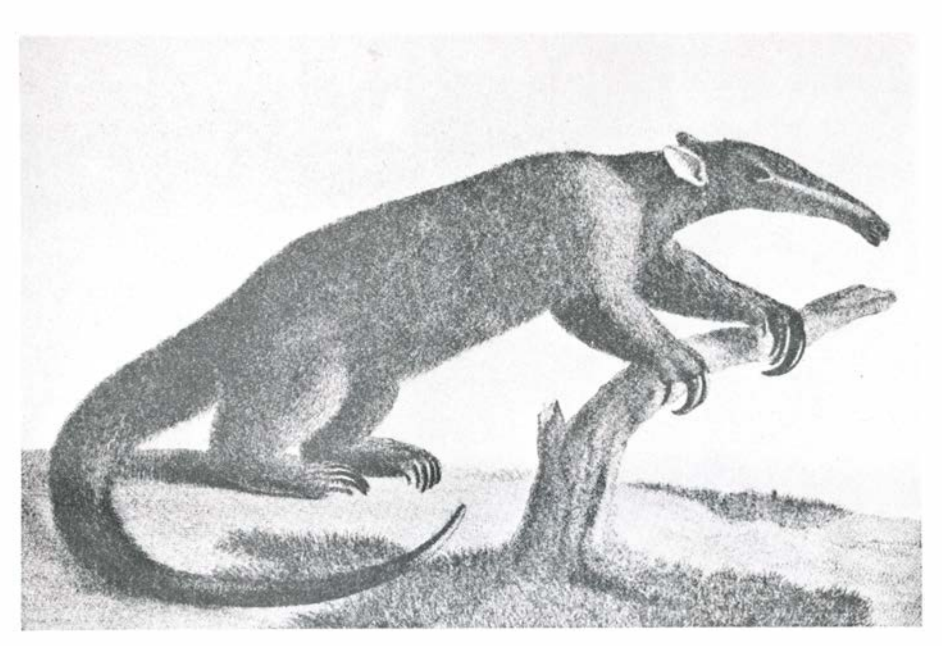
vió un ejemplar de su estudio sobre los "cuadrúpedos" a su hermano, el famoso diplomático y erudito José Nicolás de Azara, que entonces ocupaba el cargo de embajador español en París. Su intención era informarse acerca del interés de su obra, pero Moreau de Saint-Méry, un estudioso amigo de su hermano, tradujo el texto remitido al francés y lo publicó con el título de *Essais sur l'histoire naturelle des quadrupèdes de la province du Paraguay* (1801). Aunque Azara lamentó esta publicación prematura, de hecho lo dio a conocer en el ambiente científico europeo con el respaldo de varios importantes naturalistas, entre los que se encontraban el conde de Lacépède, continuador de la *Histoire naturelle* de Buffon tras la muerte de éste, y el gran Cuvier.

Tras su regreso a España, Azara se ocupó inmediatamente de publicar la versión definitiva de su obra sobre los "cuadrúpedos" y también su estudio sobre las aves. Dedicadas ambas a su hermano José Nicolás, los *Apuntamientos para la historia natural de los quadrúpedos del Paraguay y Río de la Plata*, en dos volúmenes, y los *Apuntamientos para la historia natural de los páxaros del Paraguay y Río de la Plata*, en tres, aparecieron impresos en Madrid en 1802. Este mismo año, Azara marchó a París, donde entró en relación con los principales naturalistas franceses de la época, entre ellos Cuvier y Geoffroy Saint-Hilaire. Rectificó algunos errores de su manuscrito de los *Viajes* y, por otra parte, descubrió las falsificaciones —hechas con plumas de especies diferentes— de algunos ejemplares disecados de aves sudamericanas existentes en el Museum d'Histoire Naturelle. Después de la estancia de Azara en París, se publicó una traducción francesa de sus *Viajes*, titulada *Voyages dans l'Amérique Méridionale* (1809) con notas del naturalista Charles-Athanase Walckenaer y del propio Cuvier, y seguida de la versión a dicha lengua del estudio de Azara sobre las aves. Resulta notable que, perdido el original castellano, los *Viajes* de Azara no se publicaran en nuestro idioma hasta 1846, año en el que apareció en Montevideo una traducción desde el francés realizada por Bernardino Rivadavia; y en la propia España hasta que se editó, en 1923, la traducción de Francisco de las Barras de Aragón, asimismo desde el francés. Anteriormente habían aparecido versiones al italiano, el alemán y el sueco de dicha obra, así como una traducción inglesa del libro sobre los "quadrúpedos" y un "Systematischer Index" de su estudio acerca de las aves confeccionado por un naturalista alemán.


*El ñandú o avestruz de Magallanes (***Rhea americana***), según Azara*

Tras la muerte de su hermano José Nicolás en 1804, Félix de Azara regresó a España, donde permaneció el resto de su vida. Tras rehusar el nombramiento de virrey de México, que le ofreció Godoy, aceptó el de vocal de la Junta de Fortificación de ambas Américas, que ocupó hasta su definitivo retiro en su localidad natal en 1808. Durante la guerra de Independencia participó activamente en la lucha contra los invasores, a pesar de sus excelentes relaciones con personalidades francesas y de que su hermano había sido amigo del mismo Napoleón. Cuando regresó Fernando VII, rechazó, sin duda por sus ideas liberales, una condecoración que le había concedido este monarca. En cambio, se ocupó de cuestiones técnicas y económicas relacionadas con su tierra aragonesa y, tras el advenimiento del régimen liberal, aceptó ser "Regidor Constitucional" del Ayuntamiento de Huesca, cargo para el que fue elegido en mayo de 1820. En octubre del año siguiente falleció a consecuencia de una pulmonía.

Por iniciativa de su sobrino Agustín, se publicaron en 1847 la obra de Azara titulada *Descripción e Historia del Paraguay y del Río de la Plata* y un vo-



*El oso hormiguero negro (***Tamandua nigra***), otra de las especies descritas por Azara*

lumen con memorias e informes de carácter principalmente geográfico. Posteriormente se han editado otros textos inéditos suyos, en especial en la República Argentina y en Uruguay.

La importancia de la obra de Azara reside, en primer término, en su extraordinaria aportación a la zoología descriptiva. Más de un centenar de especies de aves americanas actualmente admitidas están oficialmente basadas en sus cuidadosas descripciones y mediciones y en la detallada información que recogió acerca de su comportamiento y distribución. Lo mismo puede decirse acerca de una cifra algo inferior de mamíferos sudamericanos, entre ellos algunos con términos asociados a su nombre, como el zorro de Azara (*Dusicyon gymnocercus*) y la zarigüeya de Azara (*Didelphis azarae*).

Azara no se limitó al plano descriptivo, sino que aprovechó sus observaciones para formular un rico conjunto de ideas, hipótesis y reflexiones acerca de cuestiones biológicas fundamentales. Se interesó, en especial, por las variaciones de los animales en libertad y en domesticidad, por su distribución geográfica, por las relaciones entre presa y depredador y entre huésped y parásito, por el origen de las especies peculiares del Nuevo Mundo y por el proceso de la selección artificial.

Frente a las ideas de Buffon y otros naturalistas de la época, demostró que la influencia del clima no podía explicar las diferencias de tamaño, forma y color que presentaban animales que habitaban en sitios distintos. Las interpretó, por el contrario, como consecuencia de un mecanismo hereditario que intentó esclarecer, sobre todo en relación con los cambios sufridos por los animales domésticos que volvían a vivir en libertad, como las vacas y los caballos cimarrones, llegando, aunque de forma limitada, a formular el concepto de atavismo. También fue el primero que empleó el término "mutación" para designar la aparición súbita de caracteres singulares en individuos nacidos de padres comunes, aunque su interpretación del fenómeno no coincide, por supuesto, con la vigente desde comienzos del presente siglo.

Desde una perspectiva ecológica, Azara no solamente distinguió entre una fauna integrada por especies peculiares de América y otra formada por especies análogas a las del Viejo Mundo, sino que acertó a ver que la primera estaba en inferioridad en la lucha por la existencia y, por lo tanto, condenada a desaparecer: "Mis cuadrúpedos —dice, por ejemplo— sin análogos en el otro continente, tienen formas extrañas, de las que algunas parecen incluso monstruosas y como en una especie de contradicción o de caos. Todos los de esta clase están en un estado de torpeza casi sin medios para defenderse y algunos saben apenas mantener su existencia que se podría llamar vegetativa; razones por las cuales ellos desaparecerán de este continente después que sea poblado." Por otro lado, formuló de modo muy claro la selección artificial: "El hombre, según sus deseos, puede elegir los caballos y yeguas de cualquier raza, y lo mismo en las de otros cuadrúpedos y pájaros, y hacer que sólo estos individuos escogidos la continúen; puede también mezclar algunas razas, lo que es un modo de injertar; y por ambos medios está en su arbitrio el mejorarlas sobre lo que serán naturalmente. De la misma manera pende de la facultad humana el separar a aquellos individuos singulares que produce de cuando en cuando la naturaleza y hacer que se perpetúen formando variedades, según veremos con los toros mochos (sin cuernos). Por dichos medios, que se han practicado a veces con estudio y a veces por casualidad, se han llegado a formar algunas de las razas y variedades que vemos en los animales domésticos, y otras proceden de la unión ilícita entre animales originalmente diferentes; lo que asimismo pende del hombre, o de la domesticidad, y no de la naturaleza, que jamás las permite en su estado natural de libertad."

Anotemos, por último, que Azara defendió la hipótesis de creaciones múltiples y sucesivas para explicar los problemas que le planteaba la distribución geográfica de las especies vivientes y, en especial, el origen de algunos animales americanos. En esta cuestión, así como en la del origen de los amerindios, se apartó de las teorías tradicionales apoyadas en la interpretación literal de la Biblia.

Entre los estudiosos de la obra de Azara, Enrique Alvarez López fue el primero que puso de manifiesto la influencia que el naturalista aragonés ejerció sobre Darwin. Señaló las abundantes citas que éste le dedicó en sus libros *Voyage of the Beagle* y *Origin of Species* y anotó los principales materiales e ideas biológicas de Azara que pesaron en la edificación del evolucionismo darwinista. Recientemente, Barbara G. Beddall ha completado y matizado los resultados de Alvarez López. Por una parte, ha recogido las citas de Azara que Darwin hizo en *The Descent of Man* y, sobre todo, en *The Variation of Animals and Plants under Domestication*, así como su presencia en la biblioteca del gran sabio británico y en las notas manuscritas que forman su *Journal of Researches*. Por otra parte, ha dejado bien claro que Azara no fue un "precursor" de Darwin en el sentido de que formulara una teoría evolucionista, sino un autor que le influyó poderosamente con sus descripciones, ideas y reflexiones. Darwin leyó muy atentamente las obras de Azara en sus versiones francesas y reconoció abierta y repetidamente la deuda que tenía con él.







UN MISIL AUTOGUIADO logra un impacto directo sobre la torreta de un carro durante unas recientes maniobras en el polígono de White Sands, en Nuevo México. El misil, denominado "Copperhead", tiene un captador electro-óptico en su cabeza que detecta un punto luminoso, creado sobre el objetivo por un rayo marcador laser, generado desde una fuente remota, ya sea en tierra o en el aire. El sistema de guía del misil utiliza las señales del captador para dirigir el proyectil hacia el objetivo. El "Copperhead", diseñado para ser disparado desde un obús de 155 milímetros, entrará en servicio en el ejército norteamericano a finales del presente año. En su versión actual tiene un alcance de unos 16 kilómetros. Hay en estudio otro modelo con un radio mayor.

Armas con guía de precisión

Los misiles guiados sobre el objetivo pueden alterar el equilibrio de fuerzas en el campo de batalla. Especialmente vulnerables a tales armas "inteligentes" son los carros de combate, que algunos países están adquiriendo ya en gran número

Paul F. Walker

Cuatro meses después del final de la guerra de 1967 en el Medio Oriente, los tripulantes israelíes a bordo del destructor *Elath* divisaron, en el Mediterráneo oriental, un largo y verde misil que se les aproximaba lentamente desde la dirección de Port Said, en Egipto. El objeto, de extraña estampa, dotado de cortas alas y aletas de cola, y volando exactamente por encima de las olas, parecía fuera de control; puesto que se encaminaba al mar abierto lejos del *Elath*, la tripulación tuvo tiempo de maniobrar el buque y comenzar a hacer fuego sobre él. Repentinamente, para asombro del comandante israelí, el misil giró y se dirigió sobre el *Elath*, alcanzándolo en su centro entre una granizada de proyectiles de ametralladora. Minutos más tarde apareció un segundo misil sobre el horizonte, e hizo explosión en la sala de máquinas, inutilizando el buque. Cuando los miembros de la tripulación se apresuraban para extinguir el fuego consiguiente y socorrer a los heridos, un tercer cohete alcanzó al incendiado buque y lo hundió. Finalmente, un cuarto misil se hundió en las olas entre los supervivientes donde el *Elath* acababa de desaparecer.

Un portavoz de la armada israelí explicó al día siguiente que el *Elath* había sido alcanzado por misiles antibuque de fabricación soviética, lanzados desde una lancha rápida egipcia de 90,71 toneladas a más de 10 millas. El *Elath*, de 2267,5 toneladas, bautizado *H.M.S. Zealous* en el momento de su botadura en Inglaterra en 1944, era uno de los dos buques gemelos de guerra adquiridos de los ingleses en 1956. Junto con su compañero de serie, el *Yaffo*, había constituido una parte substancial de la armada israelí por más de una década. Sin embargo, fue destruido, pereciendo una cuarta parte de su tripulación de 202 hombres, por un buque de apenas una fracción de su tamaño y coste. Y acabó con él un misil de seis metros de longitud, con aletas en delta, propulsado por un motor-cohete y diseñado por los rusos a finales de los años 1950; conocido por los servicios de información occidentales como el "Stix", estaba dotado de un sistema de guía terminal radárica y una cabeza de guerra con 454 kilogramos de alto explosivo.

Este incidente, el primer hundimiento de un buque por un misil autoguiado, no tripulado, constituyó un hito en la guerra naval. Fue también una inicial y costosa demostración del impacto potencial sobre las operaciones militares de las armas "inteligentes": proyectiles convencionales (no nucleares) que pueden ser guiados exactamente hasta su objetivo después de su lanzamiento. Revisaremos aquí la evolución de los proyectiles con guía de precisión, su estado actual en los arsenales de las principales potencias militares y la controversia sobre su probable papel futuro, particularmente en el campo de batalla terrestre. Donde sea posible, citaré los resultados de la experiencia en combate real y en ensayos de polígono. Ha de recordarse, sin embargo, que una gran cantidad de información acerca de estas armas permanece secreta.

En el incidente del *Elath*, un buque pequeño rápido, y altamente maniobrable, armado tan sólo con dos lanzadores de misiles antibuque, se las arregló para vencer en combate con un buque de mucho mayor tonelaje, armado con 10 cañones contra superficie y antiaéreos y ocho tubos lanzatorpedos. Un poderoso Goliath, de formidables características contra cualquier enemigo dentro del alcance de sus cañones, fue vencido por un minúsculo David, que le dio muerte desde mayor alcance en un ataque por sorpresa. Este incidente, y muchos otros posteriores en tierra y mar, han confirmado el incremento no sólo del alcance, sino también de la potencia destructora, de las armas en el campo de batalla táctico. Aunque los misiles Stix, de hace una década o más, eran de burdo diseño en comparación con las armas guiadas actuales, lograron hundir un buque de guerra cuyo coste era miles de veces mayor que el suyo propio.

La relación coste-eficacia de este modo de ataque no pasó inadvertida. Las principales escuadras del mundo están realizando importantes inversiones en buques diseñados para lanzar misiles con guía de precisión. El efecto más inmediato del incidente del *Elath* se patentiza en la modernización de las armadas egipcia e israelí. Israel ha retirado del servicio el destructor superviviente, el *Yaffo*, y está construyendo una flotilla de 44 lanchas rápidas armadas con un misil antibuque con guía de precisión y fabricación israelí: el "Gabriel". El desplazamiento promedio de los buques en la flota de superficie israelí ha disminuido desde 421 toneladas en 1967 hasta 181 toneladas, aproximadamente, en la actualidad. Egipto ha adquirido 54 lanchas rápidas armadas con misiles, ha retirado del servicio el grueso de sus mayores buques de guerra y ha reducido el desplazamiento promedio de su flota de superficie desde 376,5 hasta 331 toneladas.

La aparición de las armas con guía de precisión ha influido también en los planes militares de norteamericanos y rusos, en relación, sobre todo, con los arsenales de Europa central, el potencial campo de batalla más fuertemente armado del mundo. Según William J. Perry, anterior vicesecretario de Defensa para Investigación e Ingeniería, las armas con guía de precisión están "originando una revolución de las mayores proporciones", que "cambiará el aspecto del campo de batalla". Perry señala que, aunque las fuerzas de las naciones del Pacto de Varsovia gozan de una ventaja de dos a uno en carros de combate sobre las de la OTAN (y posiblemente esta ventaja se incrementaría a tres a uno, dos semanas después

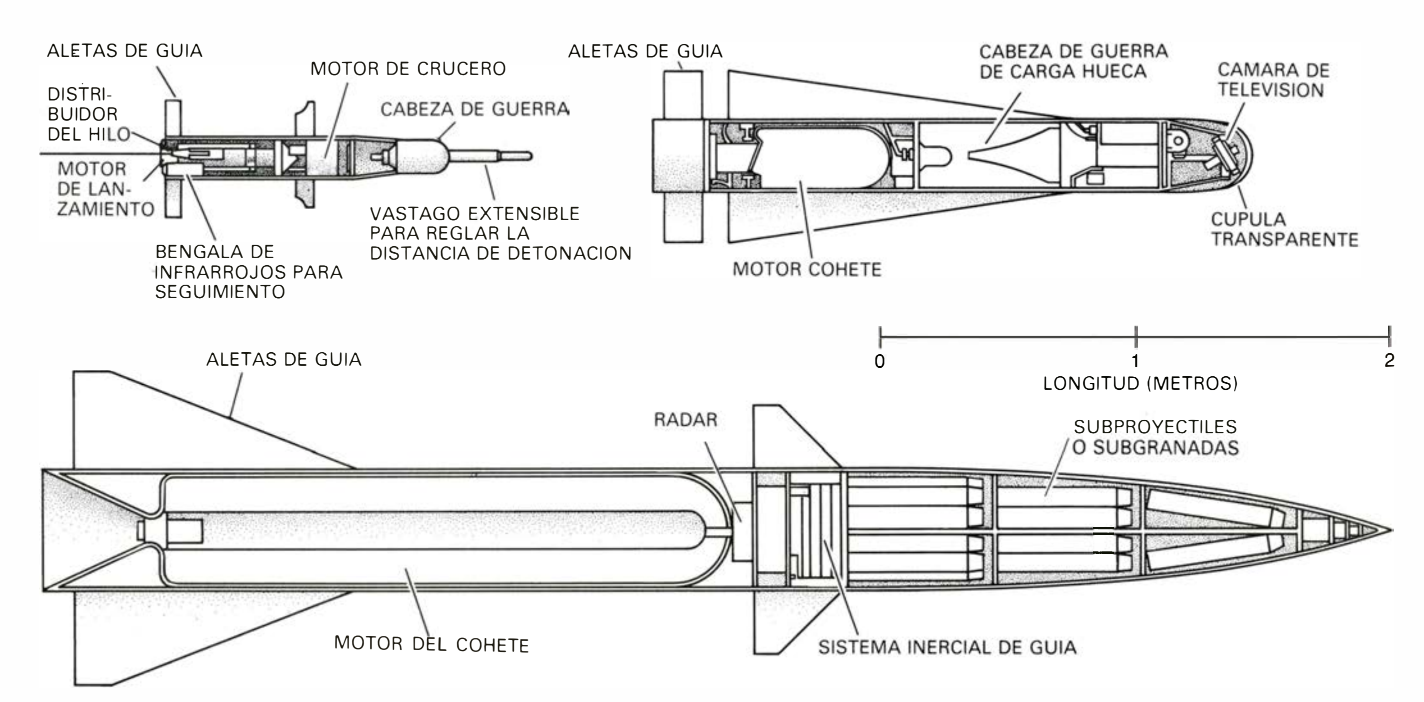
de la movilización), la adquisición de armas contracarro con guía de precisión por las fuerzas de la OTAN "debiera compensar aquella desventaja" y contribuir a mantener el equilibrio de poder en Europa central. De manera similar, el mariscal Andrei Grechko, Ministro de Defensa de la Unión Soviética, ha escrito que "los combates en el Oriente Medio... han planteado de nuevo la cuestión de la relación entre ataque y defensa... Los carros de combate son más vulnerables y su empleo en el campo de batalla, más complicado".

La idea de un misil que pudiera ser guiado a su objetivo por control remoto después de su lanzamiento se remonta largo trecho en la historia militar. Durante la primera guerra mundial, el Departamento de Guerra de Estados Unidos experimentó con una bomba volante programada llamada "Bug", que nunca se utilizó en combate. Los críticos de la nueva arma señalaron que sistemas de guía y control, de escasa confianza, podrían dañar las fuerzas propias. Esta inquietud estaba en parte justificada. En un vuelo de prueba, en 1923, de un hidroavión con un control remoto primitivo, llamado "Wild Goose" ("ganso salvaje"), el sistema de guía falló, y un "oficial de seguridad" tuvo que darle caza en un viejo biplano De Havilland. El piloto perseguidor tenía que arrojar ladrillos contra la hélice del "Wild Goose" en caso de que el avión errante amenazara estrellarse contra un área poblada.

Una versión más perfeccionada, basada en la idea de guiar el misil a lo largo de un rayo luminoso buscador, se estudió ya en 1925. El misil tenía que maniobrarse por medio de un sistema de control que incorporaba en las aletas de cola cuatro células de selenio sensibles a la luz. Sin embargo, hasta la segunda guerra mundial, ni Alemania, ni los Estados Unidos se comprometieron decisivamente en el desarrollo de misiles guiados, perfeccionados. Hacia el final del conflicto, los alemanes tenían entre manos 140 programas de misiles guiados, al menos.

Uno de los primeros misiles dirigidos estadounidenses, denominado "Weary Willie", fue un bombardero B-17, cargado con 9 toneladas de explosivos. En la operación "Kastor", puesta en práctica entre finales de 1944 y principios de 1945, aviones de este tipo participaron en seis misiones contra los refugios submarinos alemanes en Helgoland. Los pilotos se arrojaban en paracaídas a 20 millas de distancia del objetivo y los aviones eran guiados el resto del recorrido por control remoto, hasta que se estrellaban en tierra. Una bomba planeadora de 0,9 toneladas, la GB-1, fue lanzada desde un B-17 sobre la ciudad de Colonia, en 1944. El primer misil autoguiado norteamericano, una bomba planeadora con alas, llamada "The Bat" ("el murciélago"), y equipada con un sistema de guía radárico, fue lanzado contra buques japoneses hacia el final de la guerra, hundiendo en un caso un destructor a 20 millas de distancia. En esa época, los Estados Unidos desarrollaron también la bomba "Azon", ("*az*imuth *on*ly"), que podía guiarse en dirección solamente; de aquí su nombre. La bomba "Azon" fue lanzada contra diversos objetivos: esclusas del Danubio, viaductos cerca del Paso del Brennero, puentes sobre el Sena y líneas de comunicación y transporte en Burma, entre otros. La bomba "Razon" ("*r*efined *Azon*"), más perfeccionada, dirigida por radio tanto en alcance como en azimuth, nunca llegó a entrar en servicio.

Los japoneses desarrollaron sus propias y peculiares bombas guiadas —los aviones kamikaze— durante la segunda guerra mundial, si bien tuvieron un éxito limitado. En la campaña de Okinawa, en 1945, el 93 por ciento de esos aviones pilotados y cargados de explosivo fallaron su objetivo o cayeron derribados. Los técnicos militares alema-



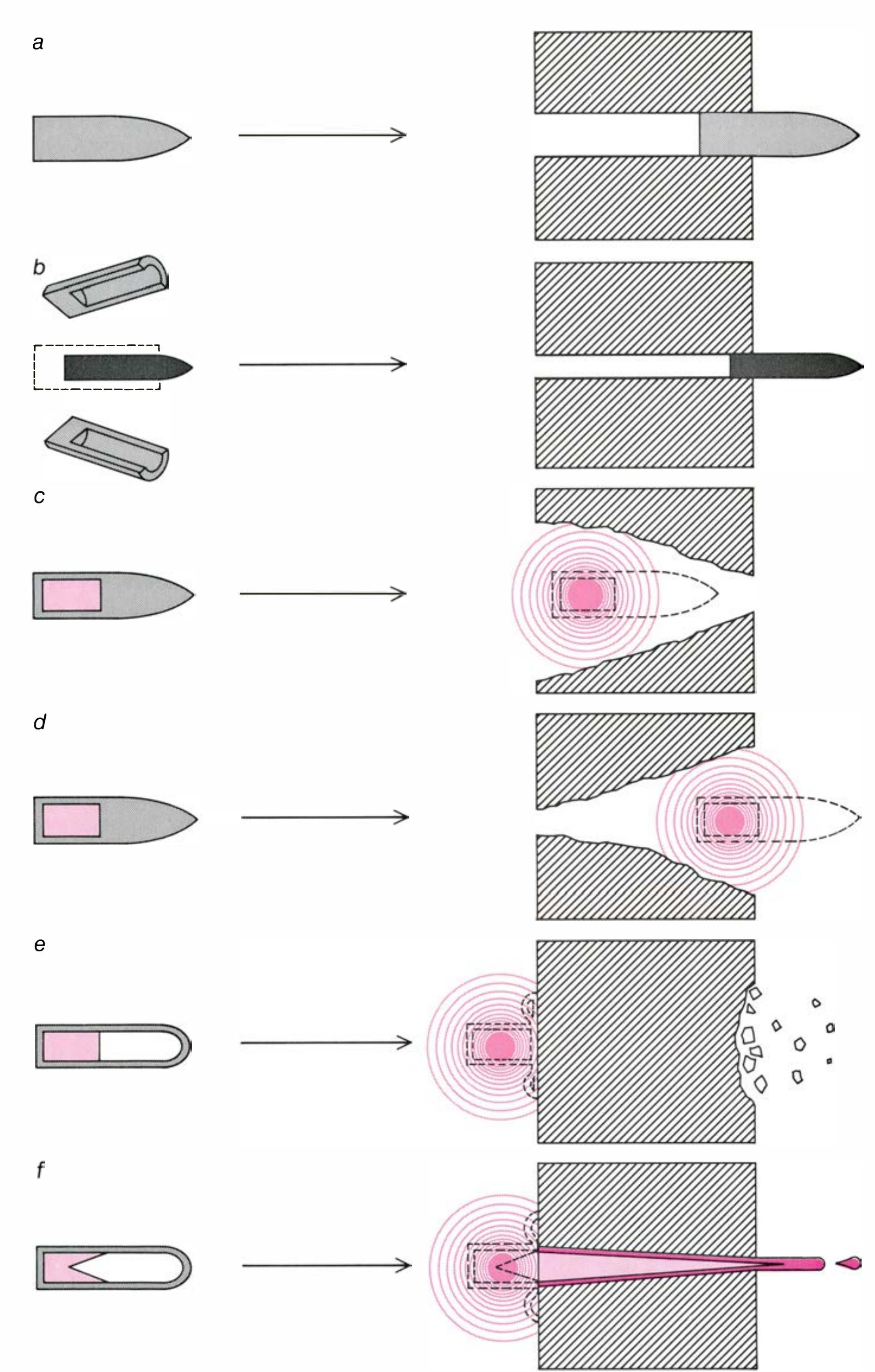
EN ESTOS CORTES LONGITUDINALES se representan tres generaciones de misiles norteamericanos con guía de precisión. La figura superior izquierda muestra una versión mejorada del misil contracarro de corto alcance "Tow", desarrollado a mediados de la década de 1960. El misil semiautomático "Tow" puede dispararse desde el suelo, desde un vehículo ligero o desde un helicóptero. Después del disparo del misil, el tirador sólo debe conservar el retículo de su visor sobre el objetivo; un captador de infrarrojos montado paralelamente a la línea de mira sigue al misil, con la ayuda de un pequeño generador de infrarrojos instalado en su cola. Las correcciones de la trayectoria del misil se transmiten al sistema de dirección por medio de un hilo de guía que se desarrolla desde su cola, a medida que el misil se aleja. En la parte superior derecha se muestra un misil aire-superficie "Maverick", de alcance intermedio, equipado con un sistema de guía basado en televisión. En este modelo, el piloto del avión atacante elige un objetivo sobre la imagen televisada y fija el misil sobre él; desde ese instante, el misil se dirige automáticamente sobre el objetivo. (El "Maverick" se produce también en otros modelos con guías laser y de infrarrojos.) Abajo, un cohete de largo alcance, que se incorporará al proyecto "Assault Breaker", con misión de ataque a las áreas de apoyo del segundo escalón de un ejército. El cohete lleva una cabeza de guerra multigranada, que puede liberar múltiples subproyectiles sobre el objetivo.

nes propusieron entretanto el cohete tripulado "Natter", lanzado verticalmente, que el piloto abandonaría en paracaídas sólo después de dirigirlo sobre el objetivo. También diseñaron uno de los primeros misiles guiados por cable, el "X-4", un misil aire-aire que podía controlarse por medio de un fino alambre, que se desarrollaba desde su sección de cola. Además, alemanes y estadounidenses experimentaron con bombas guiadas por televisión, mucho antes de que los programas de televisión se emitieran de una forma regular.

Estas primitivas armas guiadas apenas influyeron en el resultado de la segunda guerra mundial. En el transcurso de la contienda, y sólo en Europa, la aviación aliada arrojó bombas por un total de 5400 millones de libras (1 libra = 0,4536 kilogramos) durante más de cuatro millones de salidas de cazas y bombarderos. Según el "Strategic Bombing Survey", de los Estados Unidos, que evaluó la eficacia del bombardeo aéreo en el conflicto, "sólo el 20 por ciento, aproximadamente, de las bombas dirigidas a objetivos de precisión cayeron dentro del área del objetivo". (Los objetivos de precisión se definían entonces como aquellos que tenían un radio inferior a los 3,5 kilómetros, medidos desde la referencia de puntería.)

Hasta mediada la década de los 50, no se introdujeron las primeras armas con verdadera guía de precisión, consecuencia de una serie de perfeccionamientos en el diseño de pequeños motores cohete y sistemas de guía. Diez años antes del dramático estreno del "Stix", el campo de batalla había sido testigo de la introducción del misil anticarro francés "SS-10" en la guerra árabe-israelí de 1956. Guiado por alambre y controlado manualmente por el tirador, el "SS-10" demostró su eficacia contra los carros de combate más antiguos, con ligero blindaje. Su sucesor, el "SS-11", podía ser disparado desde un helicóptero.

Esta primera generación de pequeñas armas "inteligentes", contracarro, que incluía, además de los modelos franceses, el misil inglés "Vigilant" y el alemán "Cobra", permitían al tirador guiar al cohete durante todo su trayecto hasta el objetivo, un avance notable sobre los primitivos sistemas. El infante podía seguir ya visualmente el objetivo y tener en cuenta sus cambios de dirección o velocidad. La carencia de una electrónica de circuitos integrados en estas armas precoces las hacía, no obstante, escasamente dignas de confianza. Por otra parte, la precisión del arma



SEIS CLASES DE PROYECTILES PERFORANTES. El proyectil perforante tradicional (*a*), macizo, utiliza su alta velocidad, perfil exterior y dureza para atravesar la pared de un vehículo acorazado. Una variante de este diseño básico es el "APDS" ("*a*rmor-*p*iercing *d*iscarding *s*abot"), o proyectil perforante subcalibrado, que consta de un núcleo pesado, de pequeño calibre, rodeado de una envolvente mucho más ligera, que se desprende cuando el proyectil abandona el tubo del cañón (*b*). El núcleo no precisa ser disparado desde un ánima rayada y puede tener mejor perfil aerodinámico. Por otro lado, el núcleo se fabrica en un metal denso (en algún caso se ha utilizado uranio empobrecido) y adquiere la mayor parte de la energía cinética del disparo, consiguiendo una capacidad de perforación mayor que un proyectil del mismo tamaño lanzado desde un tubo de menor calibre, con la misma presión en recámara. Los proyectiles "HEAP" ("*h*igh-*e*xplosive *a*rmor *p*iercing"), perforantes de alto explosivo, poseen una cabeza de acero duro y una carga explosiva que detona por medio de una espoleta de percusión con retardo; el proyectil se empotra (*c*) en el blindaje o pasa a su través (*d*) antes de hacer explosión, logrando mayores efectos que un proyectil ordinario de alto explosivo. Los explosivos del tipo "HESH" ("*h*igh-*e*xplosive *s*quash *h*ead"), de alto explosivo y cabeza de aplastamiento (*e*), se desmoronan al impacto para lograr una gran superficie de contacto antes de la detonación del explosivo; cuando éste detona, se establece una violenta onda de choque a través del blindaje, que puede originar el descascarillado de la superficie interior, proyectando fragmentos a alta velocidad. La cabeza de guerra perforante más común en infantería es la "HEAT" ("*h*igh-*e*nergy *a*nti*t*ank"), contracarro de alta energía, que transporta una carga hueca, iniciada desde su parte posterior, en el instante en que la parte frontal, que contiene la carga explosiva con un vaciado cónico, entra en contacto con el blindaje (*f*). La onda de detonación se concentra así en un chorro de gases calientes, capaces de perforar espesos blindajes, introduciendo un chorro de metal fundido en el carro.

dependía por entero de la habilidad del operador, que debía guiarla sobre un recorrido de dos o tres kilómetros, conservando ambos, misil y objetivo, sobre la línea de mira. En un campo de batalla lleno de polvo y humo era una tarea difícil, incluso para el soldado más decidido, y podía resultar aún peor por el mal tiempo u otras perturbaciones, como el fuego de la artillería enemiga. Una pequeña bengala instalada en la cola del cohete ayudaba a mantenerlo sobre la línea de mira mientras se alejaba rápidamente, pero muchos analistas militares contemplaban escépticos la eficacia de tales armas.

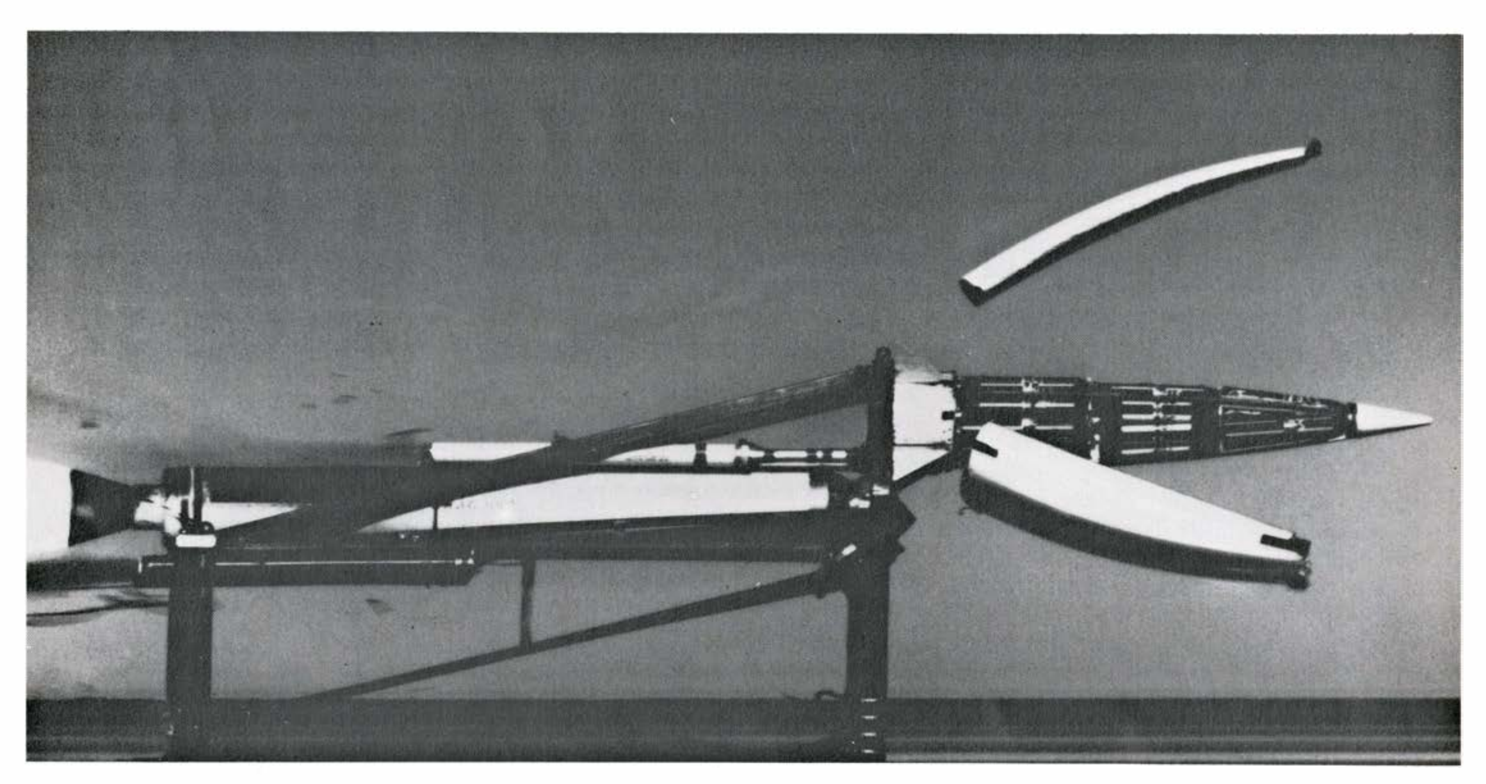
Entre los primitivos misiles contracarro, filodirigidos y manualmente guiados, uno de los primeros fue el "Sagger", fabricado en Rusia. Se trata de un arma ampliamente desplegada todavía en la Unión Soviética, Europa oriental y Oriente Medio. El "Sagger", un cohete de 11,34 kilogramos con una cabeza de guerra de alto explosivo, puede ser transportado por un infante, o montado sobre un vehículo blindado. Suelen lanzarlo un pelotón de tres soldados. Se dice que tiene un alcance de unos 3 kilómetros y que es capaz de perforar una placa de acero de 40 centímetros.

El "Sagger" apareció por primera vez en Vietnam el 23 de abril de 1972, donde destruyó varios carros de combate de fabricación norteamericana, tripulados por soldados sudvietnamitas. Un informe estadounidense sobre el incidente nos describe así el ataque: "Al principio, los carristas del ejército sudvietnamita quedaban atónitos por el lento y errático vuelo del misil. Tras breves tanteos, la tropa aprendió enseguida a atacar el punto de lanzamiento... con el fuego del cañón principal del carro, y realizar maniobras evasivas con su vehículo. (No obstante), a media mañana todos los oficiales del tercer escuadrón habían perecido o estaban heridos, y tres carros M-48A3 habían sido destruidos por los misiles "Sagger"."

La demostración de la eficacia de los misiles "Sagger", a la que suele aludirse con mayor frecuencia, ocurrió en la guerra árabe-israelí de 1973, cuando las armas contracarro se cobraron pesado tributo en ambos frentes, el sirio y el egipcio. Chaim Herzog, mayor general retirado del ejército israelí, relata el éxito del "Sagger" contra las fuerzas acorazadas hebreas en su libro *The War of Atonement*: *October 1973*, al describir los campos de batalla sembrados de carros de combate incendiados y de una especie de tela de araña formada por los hilos de los misiles guiados. En los altos del Golán se destruyeron más de 1500 carros de combate sirios, israelíes, jordanos e iraquíes, muchos de ellos por misiles contracarro. En el frente egipcio, una brigada israelí de carros perdió 77 vehículos, tres cuartas partes de su dotación al completo, en 16 horas de emboscadas contracarro. El general egipcio Sa'ad Al Shazli dedujo que "era imposible garantizar el éxito de un ataque, ya fuere de carros, ya de infantería blindada, sin destruir o silenciar antes los misiles contracarro". Herzog puntualiza que, hasta la guerra, los estrategas israelíes demostraron "una absoluta falta de apreciación de la nueva capacidad contracarro por parte de las fuerzas árabes".

Por aquel entonces, se estaba abriendo ya camino un avance importante en el armamento guiado contracarro: la introducción de la guía semiautomática. Nuevas armas tales como el misil americano "Tow" (*t*ube-launched, *o*ptically-tracked, *w*ire-guided: lanzado desde tubo, con seguimiento óptico y guiado por cable), desarrollado a mediados de los años 60, no necesitaban ser dirigidos a su objetivo mediante una palanquita accionada a mano. El sirviente tan sólo ha de retener el retículo de un visor sobre el objetivo. Un captador de rayos infrarrojos montado sobre el visor sigue al misil, dotado de una pequeña bengala de rayos infrarrojos en su cola. El visor integrado genera, a su vez, señales eléctricas, que son transmitidas a través de un hilo, corrigiendo automáticamente cualquier desvío entre la línea de mira y la traza del misil.

El misil "Tow" proporciona al solda-



ENSAYO SOBRE RIEL DE ALTA VELOCIDAD del sistema de dispersión de subgranadas para el proyecto Assault Breaker, realizado el 20 de agosto de 1980 en el Naval Weapon Center de China Lake (California). La fotografía de la izquierda muestra las envolventes exteriores de la cabeza de guerra en el

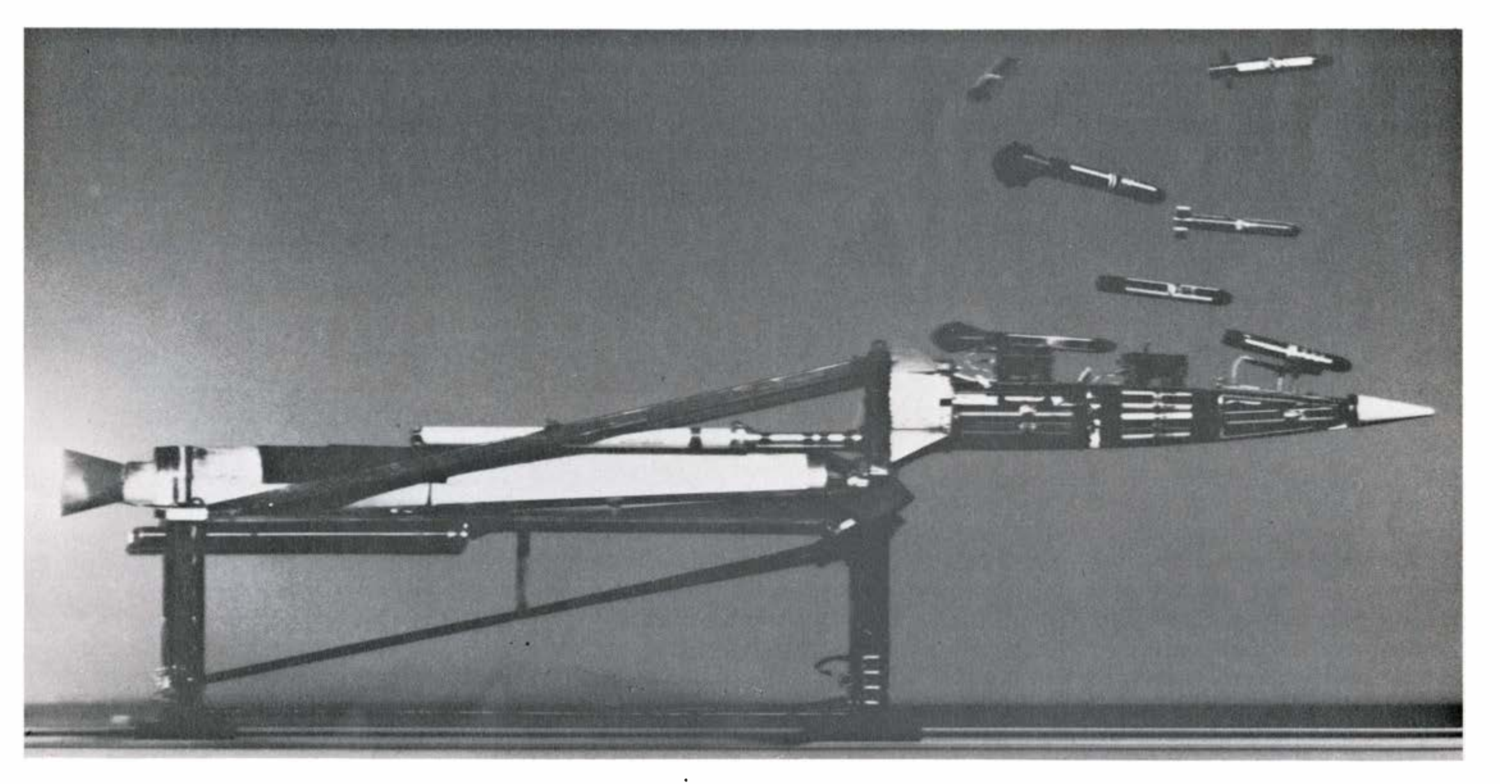
do de infantería mejores oportunidades de destruir un carro con un disparo único, particularmente a larga distancia. Ensayado en combate por primera vez en Vietnam el 24 de abril de 1972, al día siguiente de la introducción del "Sagger", el misil "Tow" se ha mostrado eficaz y digno de confianza. Más de 275.000 misiles han sido comprados por 33 países, incluida la partida de 107.000 adquirida por el ejército norteamericano y el Cuerpo de Marines. El ejército los está recibiendo a un ritmo de 12.000 por año, y los despliega tanto para ser disparados por la infantería, como para ser montados en helicópteros. Se está mejorando la calidad del misil, con una cabeza de guerra perforante más refinada y un sistema de guía más preciso, regido por un microprocesador.

Armas contracarro pequeñas, baratas y disparadas por el soldado de infantería, como el "Tow", arrastran todavía un grave inconveniente de los primitivos diseños: el hilo, "cordón umbilical" que los hace vulnerables al fuego enemigo, en particular de los cañones de alta velocidad de los carros. Si la vigilante tripulación del carro percibe el fogonazo de lanzamiento de un misil contracarro, o ve al misil en vuelo, un disparo en la dirección del punto de lanzamiento puede bastar para romper el hilo desenrollado. Inconveniente que la mayoría de los misiles modernos contracarro salvan al ir controlados por radio, operando comúnmente sobre varias frecuencias, a fin de impedir interferencias u otras contramedidas electrónicas. El misil contracarro ruso "Swatter", de 27 kilogramos, diseñado para su lanzamiento desde vehículos blindados, constituye un ejemplo característico de la actual generación de misiles controlados por radio; se supone que está desplegado en Europa oriental y Oriente Medio.

La eliminación del cable de seguimiento no resuelve enteramente el problema de la vulnerabilidad al fuego enemigo. Hasta un disparo rápido e impreciso hacia el soldado que gobierna el misil puede entorpecer su visión o distraer su atención. Para evitar esta dificultad, la próxima generación de armas contracarro será predominantemente de misiles con guía automática, del tipo "dispara y olvida" (fire and forget), en vez de las actuales versiones, del tipo "dispara y persigue" (fire and follow). Los Estados Unidos han adoptado ya sistemas de guía basados en televisión, en captadores de infrarrojos, que detectan el calor del objetivo, y en marcadores de rayos laser. En el diseño de los misiles equipados con tales sistemas se tiene en cuenta ya su origen de disparo: desde buques, aviones o helicópteros, a causa de su mayor tamaño. La eficacia de tales armas contra objetivos difíciles se demostró en 1972, cuando los Estados Unidos reanudaron el bombardeo de Vietnam del Norte, tras una prolongada moratoria. Los pilotos norteamericanos, provistos con nuevas bombas guiadas por televisión y laser, lograron destruir 106 puentes en tres meses, entre los que se contaban varios que habían resistido repetidos ataques durante la anterior fase de bombardeo sin armas inteligentes.

De entre las armas guiadas de fabricación norteamericana, una de las más ampliamente desplegada es el "Maverick", un misil aire-superficie que se presenta con tres clases de sistemas de guía a elección. En uno de ellos, una cámara de televisión montada en la cabeza del misil transmite una imagen a una pantalla en la cabina del avión atacante. El piloto selecciona un objetivo sobre la pantalla, mueve un retículo para coincidir con él y fija el misil sobre el objetivo. Desde ese instante el misil avanza automáticamente hacia el objetivo, dejando libre al piloto para dedicarse a otros cometidos, o evadirse del fuego enemigo. Muy parecido es otro de los sistemas, salvo en que utiliza un captador de rayos infrarrojos para detectar ligeras diferencias en la cantidad de calor radiado por diversos objetos. El tercer método, que es el preferido por el Cuerpo de Marines, emplea un sistema electro-óptico que sigue a un punto de luz laser reflejada desde el objetivo. El marcador laser puede accionarse desde el avión atacante, desde otro avión o desde tierra.

Esta clase de sistemas de guía se han



instante de desprenderse y liberar el sistema de dispersión. La fotografía de la derecha muestra las cinco clases de subgranadas en el instante de ser expulsadas de la cabeza de guerra. El sistema de dispersión fue construido por la firma Vought Corporation para la Advanced Research Projects Agency.

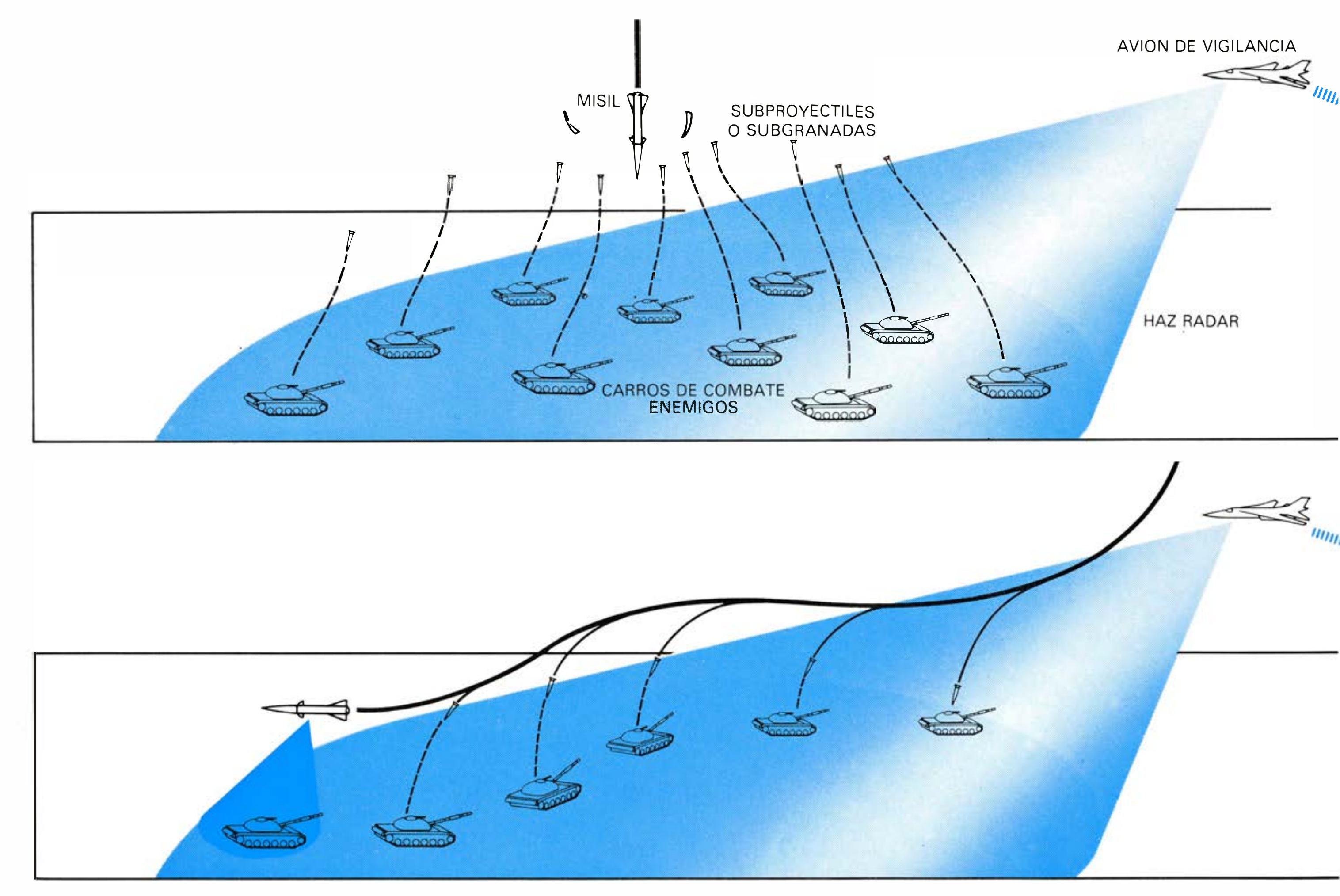
comportado bien tanto en tierra, como en el mar y el aire. Se asegura que el "Maverick" se ha apuntado un impacto directo en más del 88 por ciento de todos sus ensayos en fuego; otros sistemas han alcanzado marcas equiparables. Pero existen importantes limitaciones sobre el funcionamiento esperado de las nuevas armas inteligentes en el campo de batalla, que habrán de superarse para que tales armas revolucionen verdaderamente los combates (y la guerra) en tierra.

El primer problema, y el más obvio, es táctico: el enemigo puede todavía ganar en agilidad táctica a los misiles con guía de precisión, particularmente a los tipos más lentos de misiles contracarro. Un misil filodirigido puede tardar hasta 15 segundos en alcanzar su objetivo, en tanto que el proyectil de alta velocidad del carro puede cubrir la misma distancia en uno o dos segundos. Como parte de su formación, se instruye a las tripulaciones de los carros para que descubran las posiciones probables de los equipos contracarro y cubran tales posiciones con su propio fuego, o soliciten fuego artillero, antes de que la formación de carros se mueva dentro del alcance de los misiles.

Otra medida táctica, a disposición del comandante del carro, es el enmascaramiento, ya sea con nubes de humo para reducir la visibilidad a las fuerzas contracarro, con bengalas para desorientar los ingenios dirigidos con infrarrojos o luces brillantes que confunden a los sistemas de guía laser. La mayoría de los carros y de los transportes blindados de personal están hoy equipados con lanza-granadas fumígenas, dispuestos a esparcir una densa nube de humos de fósforo que duran varios minutos, tiempo suficiente en la mayor parte de los casos para introducirse en la zona de fuego de los emplazamientos contracarro.

Las condiciones atmosféricas pueden llegar a constituir otro impedimento para las armas con guía de precisión. Los sistemas de televisión funcionan mejor en días luminosos y con atmósfera limpia; la niebla, la lluvia y la nieve (que no son infrecuentes en Europa central) pueden alterar en grado apreciable el comportamiento de la mayoría de los sistemas de guía de esta clase. Tales consideraciones tienden a primar las versiones con guía laser o de infrarojos del "Maverick", por ejemplo, sobre el tipo con guía de televisión.

Otra contramedida disponible, en especial para los carros, consiste en incrementar su propia protección en forma de blindaje. Los carros modernos están cubiertos de pesadas placas de acero laminado y perfeccionados materiales compuestos, para protegerlos de la incrementada capacidad de perforación que poseen las nuevas armas. Sin embargo, en los últimos años, en que el peso de los carros se va acercando a 45 o 55 toneladas, el coste del blindaje añadido, particularmente en pérdida de velocidad y movilidad, ha llegado a ser prohibitivo. Parece, pues, que el blindaje está perdiendo el pulso que ha venido manteniendo con las cabezas de guerra perforantes de precisión creciente. Las nuevas cabezas de guerra, con cargas huecas de alto explosivo, al originar la detonación originan un es-



ESTAS VISTAS PANORAMICAS del campo de batalla ilustran dos aplicaciones propuestas para el Assault Breaker (el proyecto más ambicioso de todas las armas inteligentes en manos del ejército). Radares de búsqueda lateral o hacia adelante, mantenidos a distancia, y operados desde aviones, helicópteros o vehículos de control remoto, localizarían objetivos a gran distancia, tales como formaciones de carros muy adentrados aún en el territorio enemigo. Se transmitirían las coordenadas de la posición del objetivo a un centro director de ataques móvil, que dispondría el lanzamiento de un misil en la dirección

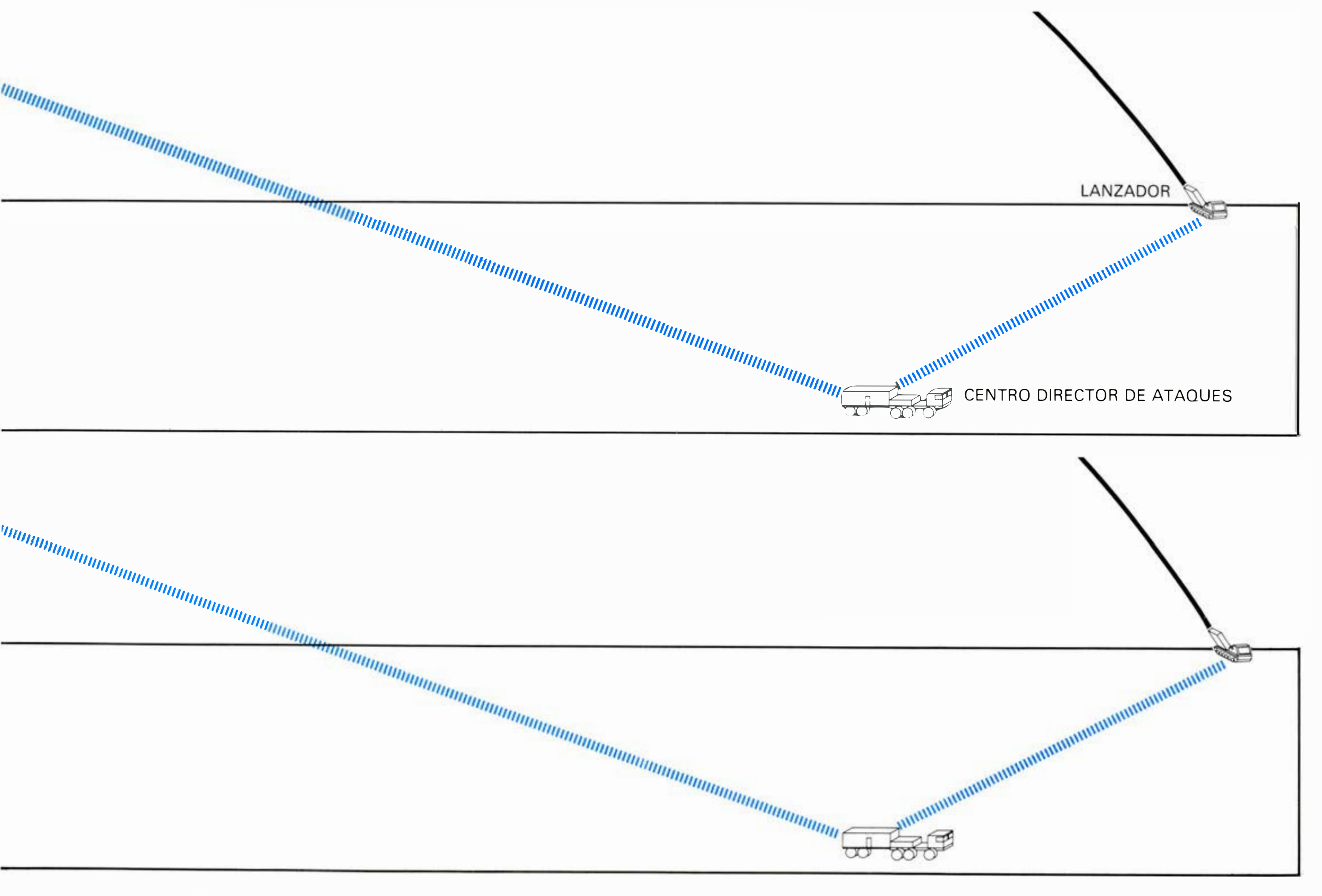
trecho chorro de gases extremadamente calientes que pueden perforar hasta cinco veces su calibre en corazas de acero; verbigracia: una granada de esta clase, de 10 centímetros de calibre, puede perforar blindajes de hasta 50 centímetros.

Durante casi dos décadas, los ingleses se han dedicado a mejorar los blindajes, sin adicionarles peso, en sus instalaciones de ingeniería y transporte militar con sede en Chobham. El 17 de junio de 1976, el secretario británico de estado para la defensa anunció un descubrimiento que, según aseguró, era el logro más importante en el diseño de carros de combate desde la segunda guerra mundial. Del nuevo blindaje, llamado "Chobham", se dice que es capaz de proporcionar una mejor protección contra todos los tipos de armas contracarro. Consiste en placas de acero y aluminio, con un entrelazado de materiales cerámicos entre las capas de metal que permite disipar la energía cinética y térmica mucho mejor que pueda hacerlo el acero por sí mismo. El blindaje "Chobham" se ha incorporado al carro inglés "Shir"; un blindaje especial similar se pretendió también en la construcción del nuevo carro principal de batalla norteamericano, el M-1, o "Abrams". Los franceses están introduciendo un blindaje que denominan "aplicado" (appliqué); consiste en placas que pueden colocarse o retirarse en función de la amenaza, proporcionando, por consiguiente, mínimo peso y máxima movilidad en cualquier circunstancia. Además, la mayoría de los nuevos carros principales de batalla tienen faldones ligeramente acorazados, para proteger las vulnerables cadenas y ruedas, y un perfil más aerodinámico, para disminuir la vulnerabilidad.

Todas estas contramedidas, consideradas hoy idóneas, para constituir una defensa razonablemente eficaz de los vehículos acorazados bajo determinadas condiciones, parecen condenadas al fracaso a la vista de los últimos avances en armas dirigidas. Los continuos progresos en precisión, fiabilidad, versatilidad, movilidad y capacidad perforante, combinados con menores costos por disparo, ponen de manifiesto que las armas con guía de precisión superarán, antes o después, la mayor parte, si no todas, las contramedidas y se irán adueñando del campo de batalla del futuro.

Entre las armas contracarro de menor alcance (menos de tres kilómetros), se cita el norteamericano "Dragon", un misil para infantería, semi-automático, filodirigido, presentado por primera vez hace cinco años; si bien está prevista su sustitución por otros sistemas nuevos. El "Dragon" ha probado su alta eficacia en ensayos de fuego contra carros pesados, pero tiene limitaciones de índole operacional: una electrónica exageradamente complicada y un mantenimiento difícil entre ellas; se le reemplazará por un misil que fía en la guía de un marcador laser, permitiendo por tanto al tirador "disparar y esconderse" con una celeridad mayor que la hasta ahora realizable.

Un segundo sucesor posible, actualmente en desarrollo, es el "Tank Breaker", un misil que se dirige automáticamente con guía de rayos infrarrojos. Un tercer programa se basa en un



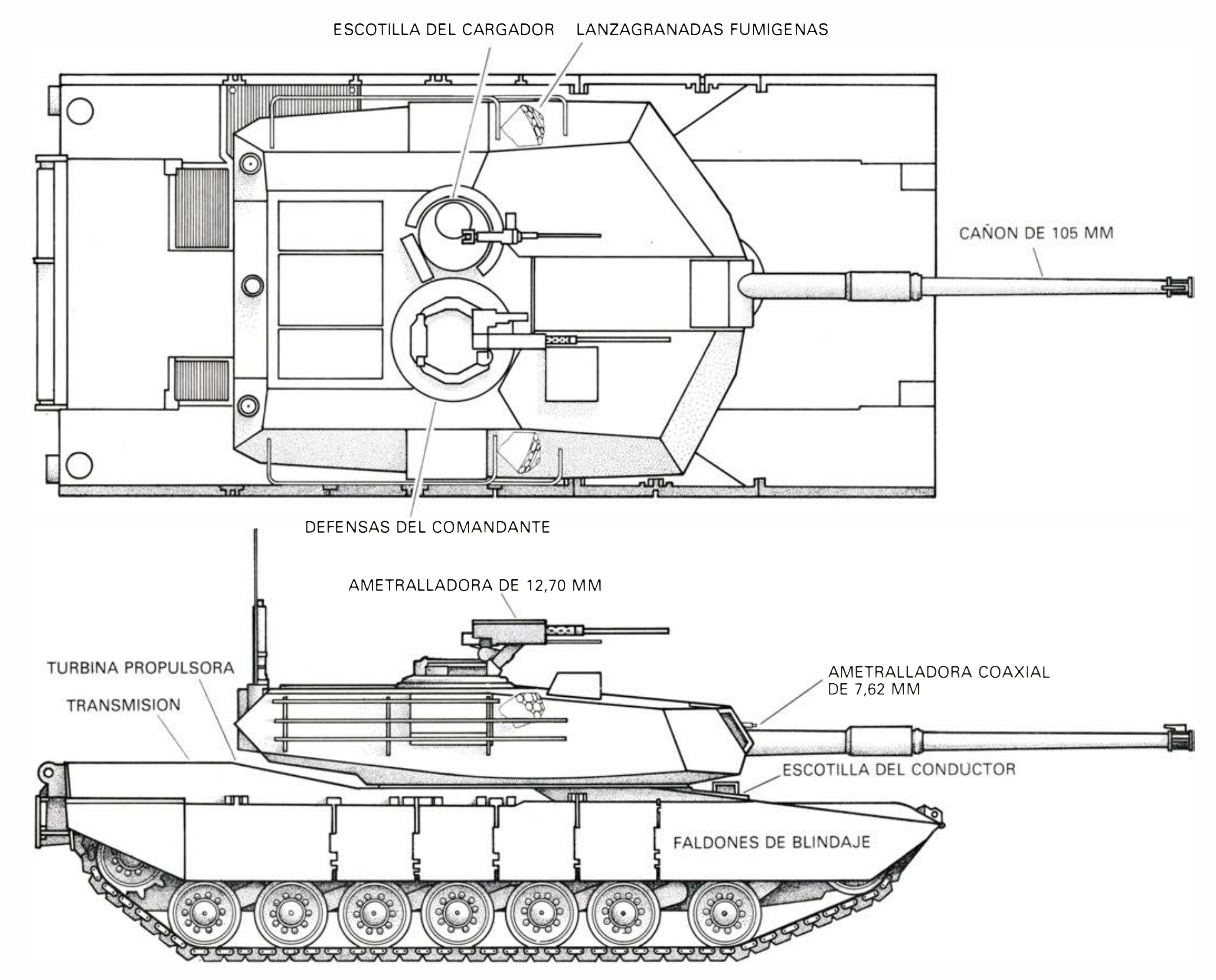
general del objetivo. Al llegar el misil sobre el objetivo se liberarían las subgranadas. Se han previsto dos modos de ataque. Para un objetivo disperso sobre un área amplia, tal como una formación de carros (*diagrama superior*), todas las granadas se liberarían simultáneamente y con sus sistemas de guía terminal, explorando con captadores de infrarrojos u ondas milimétricas, se dirigirían automáticamente sobre los distintos carros. Para un objetivo lineal, pensemos en una columna de carros (*abajo*) el misil utilizaría su radar para seguir la forma del objetivo y liberar subgranadas en la debida secuencia.

proyectil de 155 milímetros, lanzado desde un tubo y dotado de un captador de radiación de ondas milimétricas para detectar vehículos acorazados. Se apuntará el proyectil para que vuele sobre el carro a una altura máxima de 30,4 metros, y, detonando al detectar el vehículo debajo, lanzará un fragmento perforante con un ángulo de unos 90 grados hacia el interior de la torreta del carro, cuya parte superior está ligeramente blindada. Denominado usualmente "Staff" (*s*mart *t*arget-*a*ctivated *f*ire and *f*orget), el misil podrá dispararse desde el hombro del tirador o desde un vehículo y tendrá un alcance entre dos y tres kilómetros.

Un modelo avanzado del "Tow" dominará pronto los tres o cinco kilómetros de alcance, típicos de ataque sobre los carros. Habilitado para operar durante la noche, incorpora una nueva cabeza de guerra de 12,7 centímetros, para mejorar la capacidad de perforación frente a los modernos carros rusos T-72 y T-80. Hay ulteriores mejoras en proyecto; verbigracia: una cabeza de guerra de 15 centímetros y una nueva bengala de seguimiento que le dará mayor visibilidad al misil entre el humo.

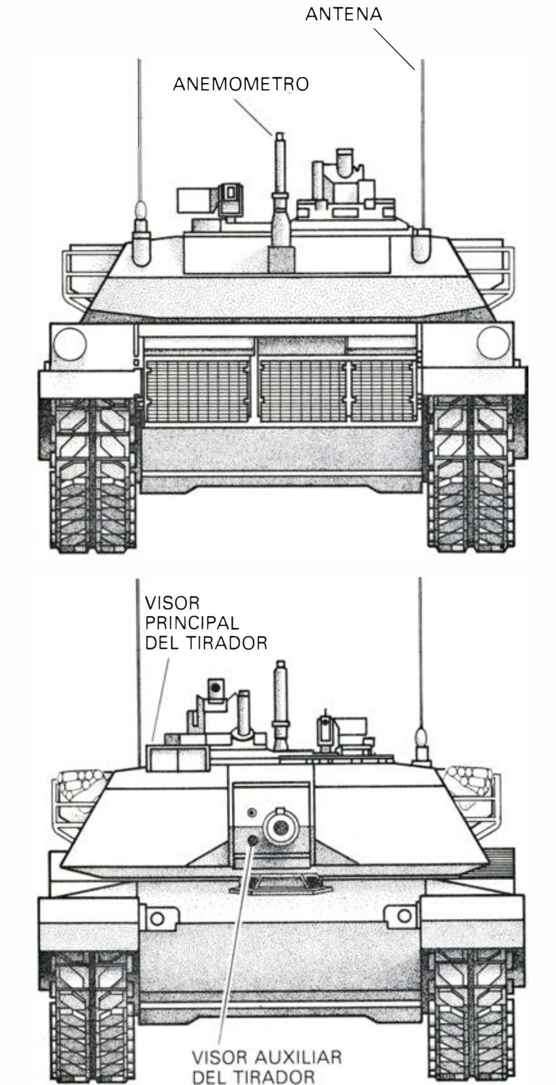
Armas inteligentes de alcance intermedio (entre ocho y 64 kilómetros) e incluso de rango mayor (hasta 200 kilómetros) habrán de aportar cambios más drásticos todavía al campo de batalla del futuro. El proyectil "Copperhead", disparado desde obuses de 155 milímetros, se destina a objetivos identificados por laser; entrará en servicio a finales de este año. El "Copperhead" ha dado ya pruebas de su contundencia en ensayos a distancia de hasta 16 kilómetros; en proyecto está dilatar su alcance hasta los 40 kilómetros. Entra en los planes del ejército de los Estados Unidos adquirir más de 2000 cohetes "Copperhead".

Un segundo sistema contracarro, de alcance intermedio, en fase de desarrollo es el "Sadarm" (*s*ense *a*nd *d*estroy *arm*or: detecta y destruye corazas). Programado para su despliegue en un plazo de cinco años, se disparará desde obuses de 20 centímetros, volará dos veces el recorrido del "Copperhead" y dispersará tres subproyectiles, cada uno de ellos dotado con un sistema miniatura de guía terminal basado en un detector pasivo de microondas. Los subproyectiles descenderán en paracaídas, explorando el suelo en busca de objetivos emisores de microondas. Según el fabricante, el captador ha sido diseñado para hacer frente "a todas las contramedidas conocidas".



NUEVO CARRO DE BATALLA norteamericano, el M-I o "Abrams". Este carro de 54 toneladas, tripulado por cuatro hombres, incorpora varias particularidades previstas para disminuir su vulnerabilidad a los misiles con guía de precisión: blindaje compuesto, un perfil más bajo y aerodinámico, faldones blindados para proteger cadenas y ruedas, y velocidad máxima de 72 kilómetros por hora. Sin embargo, se ha subido drásticamente el coste del carro durante los tres últimos años. Cifrado originalmente en unos 500.000 dólares, se ha elevado a casi tres millones: una 400 veces el valor de un misil contracarro. La Chrysler Corporation ha fabricado más de 900 carros M-I en su factoría de carros para el ejército, con sede en Lima (Ohio), por un coste estimado

El sistema MLRS, de lanzacohetes múltiple, desarrollado conjuntamente por los Estados Unidos y sus aliados de la OTAN, y cuya entrada en servicio está programada para 1982, saturará una zona con pequeños subproyectiles en un radio de unos 37 kilómetros. El proyecto permite disparar 12 cohetes en "ráfaga", o rápida descarga. Cada cohete contiene 600 granadas; se asegura que una descarga del MLRS cubre un área del tamaño de seis campos de fútbol con 7200 subproyectiles del tamaño de una granada, capaces de perforar un blindaje ligero. El ejército se propone adquirir 360.000 cohetes MLRS en los próximos años. Posteriores diseños pueden incorporar submuniciones autoguiadas y buscadores laser similares al del "Copperhead". También hay planes para cohetes todavía más refinados, capaces de fijarse sobre un objetivo después del lanzamiento y de discriminar entre carros de combate y otros vehículos o señuelos y perseguir entonces el objetivo más conveniente.



de 2100 millones de dólares. Se espera una inmediata decisión sobre duplicar o no la capacidad de producción actual del M-I, que se cifraría entonces en 60 carros por mes a lo largo del próximo año 1982. (La ilustración es de Andrew Christie.)

Pero de todas las armas inteligentes, en manos del ejército norteamericano, el proyecto más ambicioso se centra en el "Assault Breaker": un sistema de misil de largo alcance capaz de batir las áreas de apoyo del ejército enemigo. Nuevos radares de búsqueda lateral o frontal en aeroplanos, helicópteros o vehículos de control remoto localizarían las columnas de carros que se hallaran muy adentradas todavía en el territorio contrario. Los radares suministrarían las coordenadas de la posición a una batería de misiles tácticos, que lanzarían una cabeza de guerra cargada con submunición autoguiada. Los radares darían la posición en cada instante, mientras el misil vuela hacia su objetivo; en las últimas etapas del vuelo, las subgranadas tomarían a su cargo la tarea del guiado. Escudriñando el terreno con captadores de rayos infrarrojos o de ondas milimétricas, las granadas harían explosión y proyectarían una masa a alta velocidad contra la torreta del carro.

Todos esos programas y proyectos dejan entrever ciertos avances tecnológicos. En primer lugar, las cabezas de guerra perforantes, en alto grado de perfección, pueden destruir ya la mayoría de los vehículos blindados con sus tripulaciones. Tal vez la coraza "Chobhan" aporte mayor probabilidad de supervivencia frente a tales cabezas de guerra, pero transcurrirán décadas antes de que se reemplacen los actuales parques de carros de combate con blindajes de acero. Y, mientras tanto, el desarrollo de cabezas de guerra conservará, a buen seguro, su ventaja sobre las propias corazas compuestas.

En segundo lugar, las armas capaces de atacar un carro desde flancos distintos del frontal (situación en que los vehículos acorazados están más protegidos) obligará a ulteriores incrementos en el blindaje; téngase en cuenta, no obstante, que carros de 54 toneladas resultan ya prohibitivamente pesados.

En tercer lugar, el alcance de los misiles contracarro esta creciendo mucho más allá del alcance del cañón del carro; este factor, combinado con una perfeccionada vigilancia de objetivos, amplía grandemente la zona vulnerable de las fuerzas acorazadas. Los carros tendrán que confiar más en las armas de apoyo, principalmente artillería de largo alcance, si quieren intervenir con éxito.

En cuarto lugar, los progresos registrados en los sistemas de guía terminal, principalmente en captadores de rayos infrarrojos y de ondas milimétricas, elevan la fiabilidad y maniobrabilidad de los misiles, al tiempo que los convierte en menos susceptibles a las perturbaciones y más versátiles ante las condiciones cambiantes del campo de batalla. Se le está sacando partido a la tecnología de las ondas milimétricas, que utiliza una "ventana" en la atmósfera, transparente a longitudes de onda desde las microondas hasta las regiones del infrarrojo lejano (es decir, entre 30 y 300 gigahertz), para disminuir la atenuación por la calima, niebla, nubes o lluvia. Los nuevos componentes de onda milimétrica, al ofrecer pequeño tamaño, alta capacidad de transporte de señales y haces estrechos, capacitarán a las fuerzas contracarro para lograr impactos directos en cualquier tiempo y en áreas remotas.

En quinto y último lugar, dado que los misiles se subdividen en docenas de minimisiles o subgranadas, cada una con su propio sistema de guía para caer en su blanco, bajará el coste por disparo contracarro. El precio de un misil contracarro es ya mucho menor que el de un carro, disparidad que se irá ensanchando con el tiempo.

¿Cómo se reflejará todo esto en una hipotética guerra en tierra? Parecen evidentes algunas secuelas. En la alternativa entre protección y movilidad, en el futuro, el énfasis se pondrá seguramente más en la movilidad que en el blindaje. En frentes fuertemente armados, como Europa central y Oriente Medio, el bando que ataque primero, al dejar en evidencia su posición, será el más vulnerable. Todo lo que pueda detectarse, ya por medios ópticos en un día claro o por radar y guía autodirigida en un campo con humo y lluvioso o durante la noche, será alcanzado y, probablemente, puesto fuera de combate.

Esta evolución no sorprenderá a los estrategas militares que hayan participado en maniobras de ataque por sorpresa durante algún tiempo, y ciertamente desde las dos últimas guerras en el Oriente Medio. Durante unas maniobras realizadas en Alemania, en 1972, los helicópteros Cobra norteamericanos, armados con misiles "Tow", "destruyeron" un promedio de 18 carros alemanes "Leopard" por cada helicóptero "perdido". Tres años más tarde, el "Combat Development Experimentation Command" del ejército norteamericano llevó a cabo ensayos de visibilidad de los carros en función de su

despliegue. El ejército averiguó que podía reducirse la vulnerabilidad de una columna de carros, hasta en un tercio, si se secundaba el perfil del terreno, en lugar de mover las formaciones según los usuales manuales de campaña.

La operación "Bright Star", el ejercicio que la nueva fuerza norteamericana de despliegue rápido realizó el pasado diciembre en el desierto egipcio, demostró, entre otras cosas, que la movilidad constituía la clave de la supervivencia en las condiciones previsibles en un campo de batalla moderno. Las tripulaciones de los helicópteros contracarro informaron que, volando bajo, rozando el contorno y los repliegues del arenoso y estéril terreno, "podíamos disparar a 3000 metros, retroceder luego y disparar de nuevo. Podríamos zamparnos a los carros enemigos mientras hubiera espacio para movernos".

En resumen, parece que los días de la guerra relámpago ("blitz-krieg") han pasado. Podría argumentarse que unidades adecuadamente combinadas, en las que participara la infantería, la artillería y unidades acorazadas, pueden prevalecer todavía frente a las defensas contracarro. Pero los datos técnicos no abonan ese punto de vista. Hasta en un campo de batalla fangoso y caótico, las armas contracarro resultarían (si no lo son ya) harto eficaces en número elevado para que puedan destrozarlas las barreras de artillería o las embestidas de los carros. Los ataques en plan "guerra relámpago" serán suicidas. Por lo que las principales posiciones defensivas se estabilizarán en todo el mundo.

La importancia de la movilidad no sólo se predica a las tropas del campo de batalla, sino que vale también para las del segundo escalón de suministros y logística. Ante el creciente alcance de los misiles de alta precisión, la dispersión de tropas, vehículos, aviones, buques y suministros se convertirá en una imperiosa necesidad para disuadir de los ataques preventivos con armas inteligentes. Si los almacenamientos instalados en el radio de alcance de las armas de un enemigo no están muy dispersos y de algún modo protegidos, el ataque preventivo podría ser una opción posible en caso de crisis. Habrá que prestar algún interés a la dirección, control y transmisiones para garantizar el seguimiento continuo de la acción del enemigo y la adecuada coordinación de las fuerzas.

Más de un comandante habrá que crea todavía que la mejor arma contracarro es otro carro. Sin embargo, los pesados modelos actuales, que rozan las 54 toneladas y avanzan a una velocidad máxima de 80 kilómetros por hora, se están viendo sobrepasados en alcance y depreciados por los misiles de precisión, con base en tierra o en el aire. El futuro de la batalla terrestre no parece pertenecer a los vehículos acorazados, pesados, costosos y fáciles de detectar, sino a los vehículos ligeros, tipo "jeep", o a vehículos de reconocimiento móviles, baratos y por consiguiente más numerosos, armados como misiles.

Los mismos avances que han puesto al descubierto los puntos débiles de los pesados vehículos acorazados, en el campo de batalla terrestre, pueden aplicarse a la guerra en el mar y en el aire. Los misiles antiaéreos y contrabuques acabarán por ocupar muy pronto el primer plano, de suerte que las fuerzas de intervención de los portaaviones o de grupos anfibios de desembarco serán obligados a permanecer lejos del frente, fuera del alcance del ataque de los misiles. Del desamparo, cada vez mayor, de la aviación nos hablan numerosos incidentes. En Vietnam, entre 1965 y 1973, la fuerza aérea norteamericana derribó 137 cazas MIG, 90 de los cuales fueron destruidos por misiles inteligentes. En esa misma zona del sudeste asiático se demostró que los helicópteros resultaban muy vulnerables incluso al fuego de las armas ligeras. Se ha informado de que, en los recientes combates entre Irán e Irak, los aviones de caza fueron repelidos con eficacia por los antiguos cañones antiaéreos dirigidos por radar, tal como lo fueron en la guerra de 1973, en el Oriente Medio. También los misiles superficie-aire se cobraron un grave tributo de la aviación israelí en 1973, derribando unos 40 aviones sobre el canal de Suez en los dos primeros días de enfrentamiento.

En estrecha dependencia con la evolución apuntada en la táctica de poner mayor énfasis en la movilidad, habrá cambios en la política de adquisiciones. En la medida en que el campo de batalla se torne cada vez más letal, crecerá la relación coste-eficacia de los equipos, relativamente económicos y fiables, que puedan ser suministrados en gran cantidad y rápidamente reemplazados. Las plataformas grandes y costosas que transporten armas de gran valor (y personal) resultarán más ruinosas con el tiempo. No quiere decir esto que las armas inteligentes sean baratas. Un misil contra carro "Tow" viene a costar unos 7000 dólares, un misil superficie-aire "Stinger" algo así como 50.000 y el misil antibuque "Harpoon" vale más de 700.000 dólares. En cada caso, sin embargo, el objetivo atacado por el misil supera, en muchas veces, el valor de éste. El carro de combate M-1, al incluir el nuevo blindaje compuesto, ha visto elevar su coste a casi tres millones de dólares, más de 400 veces el precio de un misil "Tow". El avión de apoyo inmediato A-10 tiene un precio de ocho millones de dólares, es decir, 160 veces el coste de un misil "Stinger"; un portaaeronaves nuclear vale más de 2000 millones de dólares, 3000 veces el coste de un misil "Harpoon".

Quienes proponen continuadas y gravosas adquisiciones de carros, aviones y buques arguyen que se necesitaron varios misiles "Stix" para hundir el *Elath*, que cada misil requiere su plataforma de lanzamiento, a sumar a su coste total, y que las grandes plataformas de fuego tales como un carro proporcionan al tirador 50 disparos, y aún más. Pero, con todo, la disparidad de costes entre estos valiosos equipos, que se han enseñoreado durante años de los presupuestos militares, y las nuevas armas con guía de precisión resulta demasiado grande para que se la ignore. En los tres últimos ejercicios fiscales, el ejército compró más de 900 carros M-1, por un coste total estimado de 2100 millones de dólares; la misma cantidad hubiere servido para adquirir más de 300.000 misiles contracarro. Los programas actuales prevén que la Chrysler Corporation aumente su capacidad de producción desde 30 carros M-1 por mes hasta 90, con unas adquisiciones mínimas de 7058 carros. Simultáneamente se comprarán 3596 carros M60A3, el modelo predecesor del M-1.

No discuto que no deba incluirse ningún carro en los presupuestos futuros, pero la adquisición actualmente prevista de otros 9500 carros principales de batalla es injustificable a la vista de las posibilidades de la nueva tecnología de las armas inteligentes. Una pequeña fracción de las inversiones previstas en carros de combate podría servir para comprar una defensa mucho más eficaz para el frente de Europa central, en la forma de un vasto número de armas contracarro autoguiadas. Si se contrasta la política de adquisiciones con los avances tecnológicos, debieran solicitarse menos carros de combate, sustituyéndolos por un número mayor de armas contracarro inteligentes. Y no será ya necesario ni deseable competir con los rusos, carro contra carro, ya sea en producción ya en su despliegue en Europa.

A medida que los pequeños misiles

consigan mayor nivel de precisión y eficacia letal, los principales componentes del equipo de combate (carros, aviones de apoyo y grandes buques de superficie) acabarán por quedar anticuados para casi todas las tareas de la batalla. Con el tiempo se les relegará a un papel ceremonial y se los exhibirá el día Primero de Mayo en Rusia, o en la celebración del Día de Acción de Gracias en Norteamérica.

Desde el punto de vista del control de armamentos, las armas inteligentes pueden convertirse en una dudosa bendición. El campo de batalla será probablemente más estático; ninguno de los bandos se atreverá a tomar la iniciativa, cuando ser visible signifique ponerse a tiro. Las fronteras políticas, como las que parten Europa central, entre la OTAN y los países del Pacto de Varsovia, pueden consolidarse hasta el punto de erigirse en fronteras militares fijas. No significa esto que retornará la guerra de trincheras del primer conflicto mundial, puesto que una defensa estática es tan vulnerable como una ofensiva patente; ambos bandos reforzarán las fuerzas defensiva móviles. Tal situación, de "tablas" de facto en el campo de batalla, puede incitar al control negociado de armamentos y al desarme en armas convencionales.

Sin embargo, con un armamento que tiende a menguar de tamaño y crecer en eficacia, la lucha puede desplazarse desde los frentes militares hasta las ciudades, las guerrillas y las actividades terroristas. En la última década hubo un promedio de dos acciones terroristas por día en algún lugar del mundo. Los terroristas saben ya hacer uso de las bombas de mano, bombas de relojería y, en algunos casos recientes, de cohetes antiaéreos. Los pequeños misiles con guía de precisión, transportados en una maleta o en cualquier turismo, dotados de capacidad para perforarlo todo, salvo los blindajes más pesados, pueden convertirse en el arma preferida por los terroristas de las décadas de 1980 y 1990. Si pensamos, por ejemplo, en el ataque a un avión de pasajeros o a una central nuclear, los proyectiles con guía de precisión pueden originar daños desastrosos. El impedir el mal uso de armas tan contundentes exigirá, claramente, una distribución restrictiva y de alta seguridad, quizá mediante procedimientos similares a los vigentes para las armas nucleares tácticas; mas, por otro lado, y al objeto de lograr su máxima eficacia, las armas deberán estar ampliamente dispersas. Habrá que encontrar algún punto de equilibrio.

Las estrellas más recientes de Orión

En esa constelación hay nubes de polvo y gas que emiten radiación a longitudes de onda del infrarrojo y de radio. Algunas de las nubes más densas parecen esconder nuevas estrellas que están expulsando chorros de gas a gran velocidad

Gareth Wynn-Williams

Cuenta la leyenda que las estrellas que delinean la magnífica constelación invernal de Orión las puso allí Artemisa, diosa de la caza, para conmemorar la muerte de su compañero Orión. Pero los astrónomos de hoy opinan que Orión simboliza el nacimiento, y no la muerte; la constelación encierra una de las más ricas y próximas salas de maternidad estelares de nuestra galaxia. Dentro de un volumen cuyo diámetro mide algunos años-luz, y a 1600 años-luz de distancia, frías y vastas nubes de hidrógeno están lentamente contrayéndose, bajo la fuerza de su propia gravedad, para dar lugar a estrellas nuevas (y, presumiblemente, a nuevos sistemas planetarios). A simple vista, pueden distinguirse algunas estrellas que se formaron en un pasado relativamente reciente y muchas más estuvieron, desde hace largo tiempo, al alcance del telescopio óptico. En los últimos años, sin embargo, los avances registrados en la detección de radiación a longitudes de onda del infrarrojo y milimétricas han permitido a los astrónomos estudiar las propias nubes gaseosas, así como las más jóvenes de entre las estrellas que se albergan en sus profundidades.

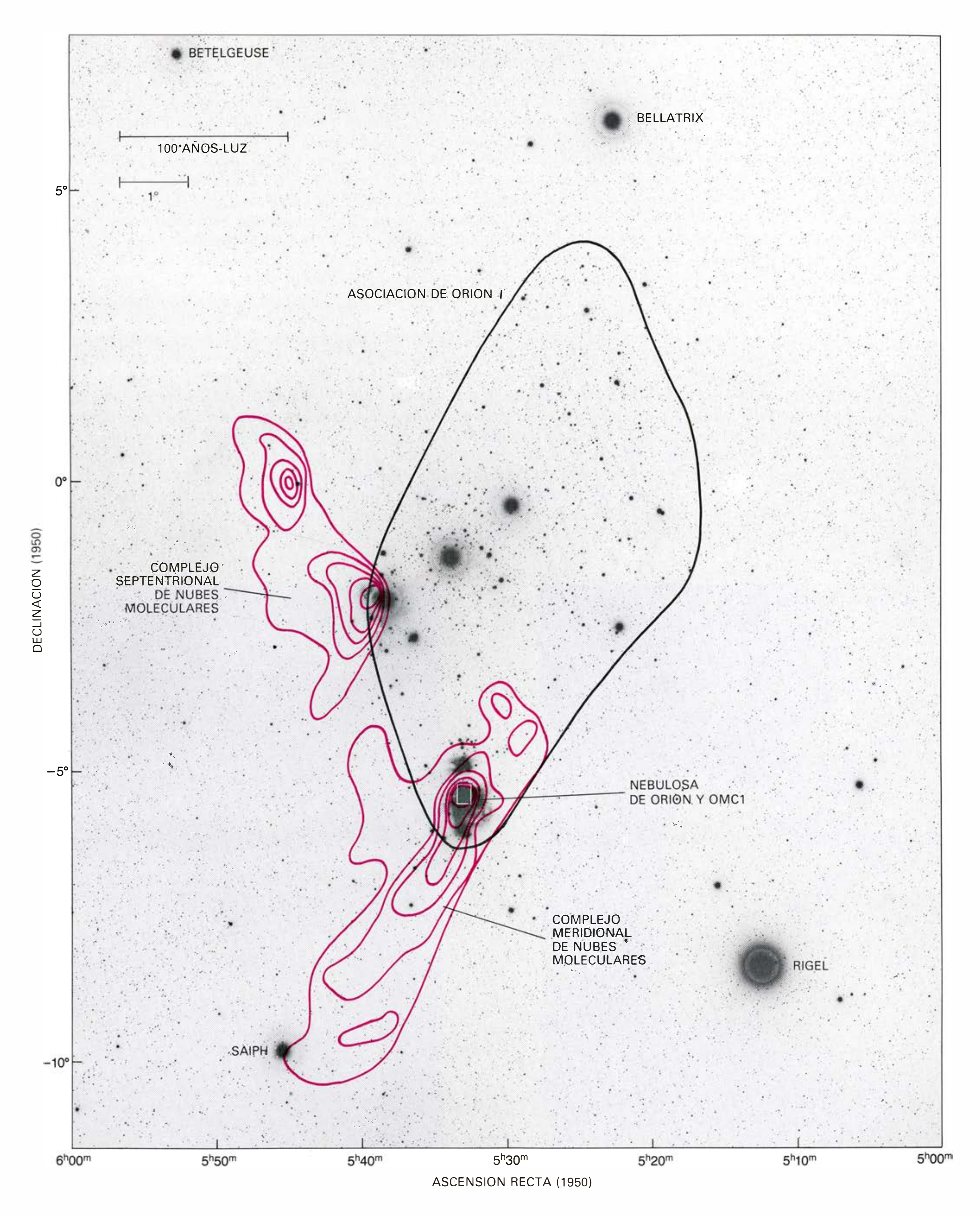
En algunos de esos jóvenes objetos, la fusión termonuclear pudo ya haber comenzado en el centro, aun cuando el espeso capullo de materia que los rodea esté todavía radiando a temperaturas altas en el infrarrojo. Mas, ¿por qué algunos objetos de esos parecen expulsar materia en vez de ir acumulándola de su entorno inmediato, como sería de esperar? Las pruebas que de ello llegan intrigan a los astrónomos.

Una estrella de la "banda principal" (una estrella ordinaria situada en la banda principal de estrellas en un diagrama que representa la temperatura en función de la luminosidad) brilla merced a la conversión termonuclear de hidrógeno en helio, en su núcleo. Cuanto mayor la masa de la estrella, más rápidamente consume su provisión de combustible nuclear; las estrellas mayores, por tanto, tienen vidas más cortas. Así, las estrellas cuya masa sea 15 veces o más la del Sol brillarán, por lo menos, 20.000 veces más que éste y convertirán todo su hidrógeno disponible en helio en menos de 10 millones de años, lo que es una pequeña fracción de la edad estimada de la galaxia: 15.000 millones de años. Las estrellas de tan gran masa, llamadas de tipo *0*, y las estrellas de masa ligeramente menor, de tipo *B*, se reconocen fácilmente por su brillo, por sus líneas características en el espectro visual y por su color azulado, indicador de temperaturas en superficie superiores a 10.000 grados Kelvin. A causa de la brevísima duración de su vida, las estrellas de tipo *0* y de tipo *B* apenas si pueden alejarse mucho de su lugar de nacimiento. De aquí que, cuando tales estrellas se encuentran en un cúmulo, se puede estar seguro de que todas ellas se formaron en fecha reciente y en respuesta a las condiciones reinantes en aquella pequeña región de la galaxia.

La constelación de Orión abunda en estrellas de los tipos *0* y *B*. Un grupo particularmente prominente es la Asociación Orión I, que abarca las tres estrellas brillantes del cinturón de Orión y las que dibujan su espada. Las estrellas de la asociación se han formado en varias épocas, a lo largo del intervalo que cubre los últimos 10 millones de años. Las más jóvenes, que probablemente no alcanzan el millón de años, se localizan en el interior y alrededores de la gloriosa Nebulosa de Orión, nube luminosa de gas ionizado que calientan un grupo de cuatro estrellas de tipo *0* conocidas colectivamente como Cúmulo del Trapecio. La estrecha relación entre jóvenes estrellas calientes y nubes de gas visibles constituye una prueba vigorosa en favor de la hipótesis de que las nuevas estrellas se forman por condensación del gas interestelar.

Para que una nube interestelar sufra el colapso gravitatorio es necesario que la presión ejercida por la gravitación mutua de sus constituyentes supere la presión térmica, que tendería a dilatarla. La temperatura de la nube ionizada en la Nebulosa de Orión es tan alta (unos 10.000 grados Kelvin), sin embargo, que el gas debería estar dilatándose, en vez de ir contrayéndose. Resulta improbable, pues, que, en el interior de la nebulosa, se estén gestando nuevas estrellas. Para buscar señales de formación de estrellas en Orión, habrá que mirar hacia nubes más frías.

Alrededor del 90 por ciento de los átomos del espacio interestelar son de hidrógeno; de helio, el grueso de los restantes. Bajo las condiciones que favorecen el colapso gravitatorio de una nube interestelar (temperaturas por debajo de 100 grados Kelvin y densidades superiores a 1000 átomos por centímetro cúbico), el hidrógeno existe principalmente en la forma de moléculas de hidrógeno, más bien que como átomos aislados de hidrógeno o iones hidrógeno (esto es, protones). Por varias razones es difícil detectar el hidrógeno molecular. Está demasiado frío para radiar a longitudes de onda del visible y, a diferencia del hidrógeno atómico, no emite radiación característica a longitudes de onda específicas que puedan detectar los radiotelescopios. Además, las nubes de hidrógeno molecular suelen ser casi completamente opacas a la luz por la presencia, en ellas, de pequeñas partículas de polvo que absorben y difunden la radiación electromagnética de longitud de onda corta. Sólo la radiación de longitud de onda considera-



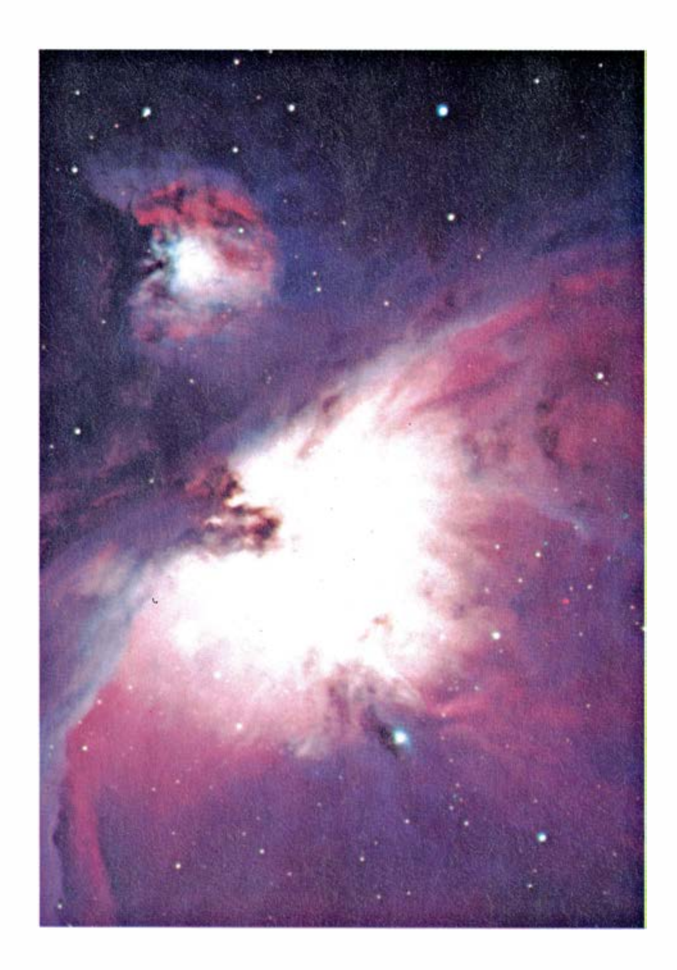
CONSTELACION DE ORION, recogida en este negativo; está espolvoreada de estrellas que son calientes, luminosas y jóvenes. Muchas de esas estrellas de tipo *0* y de tipo *B*, cuya edad no llega a los 10 millones de años, están reunidas en la Asociación Orión I. Las más jóvenes de entre ellas se encuentran en la proximidad de la Nebulosa de Orión, distante unos 1600 millones de años-luz del sistema solar. Las isolíneas perfilan la distribución de intensidad de la radiación a la longitud de onda de 2,6 milímetros procedente del monóxido de carbono, presente en pequeñas cantidades en nubes de gas molecular formadas principalmente por hidrógeno. A la baja temperatura de las nubes (menos de 20 grados Kelvin) no hay emisión detectable por parte del hidrógeno. La emisión del monóxido de carbono se observó con un radiotelescopio de 1,2 metros en la Universidad de Columbia por Patrick Thaddeus y sus colaboradores del Instituto Goddard de Estudios Espaciales. La concentración de gas, que aparentemente coincide con la Nebulosa de Orión pero en realidad está detrás de ella, se llama Nube Molecular de Orión 1 (OMC1). Es parte del complejo meridional de nubes moleculares. El rectángulo blanco delinea el área comprendida en la fotografía de la página siguiente. En esta fotografía y en otras que se sucederán, el Norte está arriba y el Este a la izquierda.

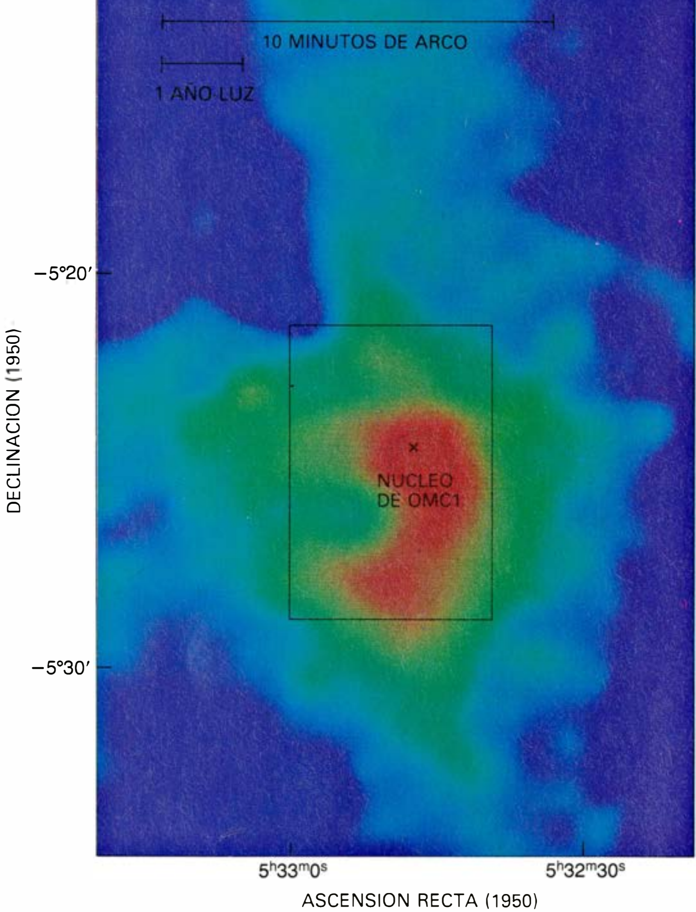
blemente mayor que la de la luz penetra las nubes moleculares o escapa de ellas. Por fortuna, hay varios procesos que generan radiación de gran longitud de onda dentro de las nubes y las hacen visibles, por tanto, a los instrumentos modernos. Los gránulos de polvo calientes emiten radiación térmica de banda ancha en la parte infrarroja del espectro, y varias moléculas distintas de las del hidrógeno emiten radiación a longitudes de onda específicas (líneas espectrales de emisión), principalmente a longitudes de onda de algunos milímetros.

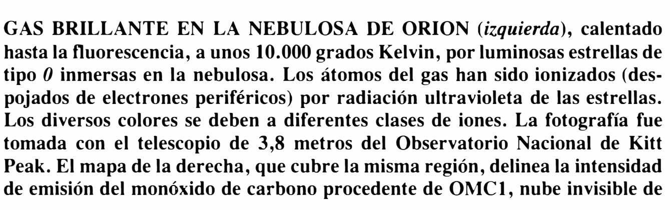
De las aproximadamente 40 moléculas interestelares observadas hasta ahora, la más útil, con mucha diferencia, es el monóxido de carbono, que detectaron en el espacio, por primera vez, en 1970, Robert W. Wilson, Keith B. Jefferts y Arno A. Penzias, de los Bell Telephone Laboratories, con un receptor sensible sintonizado a la longitud de onda de 2,6 milímetros. Los niveles energéticos responsables de la emisión de radiación electromagnética por una molécula sencilla, tal como la de hidrógeno o monóxido de carbono, dependen de la estructura electrónica de la molécula, de la vibración de los átomos de la molécula a lo largo del eje de la misma y de la rotación de la molécula en torno a un eje perpendicular al suyo. En una nube molecular, las moléculas se hallan en el estado electrónico de energía más baja, de modo que las únicas variaciones que hay que tomar en consideración son las debidas a vibración y rotación.

En general, las variaciones ocurridas en la energía de vibración de una molécula son mucho más pronunciadas que las variaciones en la energía de rotación. De donde conviene considerar cada estado de vibración como subdividido según la energía de rotación de la molécula [*véase la ilustración de la página 29*]. Un nivel dado puede caracterizarse entonces por dos números cuánticos: v para el estado de vibración y J para el de rotación. Estos números cuánticos pueden tener cualquier valor entero igual o mayor que cero. El que una molécula pueda o no absorber o emitir un cuanto de energía, y realizar una transición de un nivel a otro, depende de reglas especiales que difieren para el hidrógeno (H_2) y el monóxido de carbono (CO). Es raro, sin embargo, que v o J cambien en más de 2 en una sola transición.

La energía, E, de cada nivel, por encima del nivel fundamental (el más bajo), se expresa habitualmente en números de ondas, calculados dividiendo la energía E entre el producto de la velocidad de la luz c por la constante de Planck $\hbar$: número de ondas = $E/\hbar c$. La utilidad de esta unidad estriba en que la longitud de onda del fotón (cuanto de radiación), emitido cuando una molécula pasa de un estado energético a otro, viene dada por el inverso de la diferencia entre los dos niveles de energía medida en números de ondas. Por ejemplo, la diferencia entre los núme-






GAS BRILLANTE EN LA NEBULOSA DE ORION (*izquierda*), calentado hasta la fluorescencia, a unos 10.000 grados Kelvin, por luminosas estrellas de tipo *0* inmersas en la nebulosa. Los átomos del gas han sido ionizados (despojados de electrones periféricos) por radiación ultravioleta de las estrellas. Los diversos colores se deben a diferentes clases de iones. La fotografía fue tomada con el telescopio de 3,8 metros del Observatorio Nacional de Kitt Peak. El mapa de la derecha, que cubre la misma región, delinea la intensidad de emisión del monóxido de carbono procedente de OMC1, nube invisible de gas molecular caliente que se halla detrás de la nebulosa. Las regiones rojas corresponden a la emisión más intensa; las azules, a la menos intensa. El mapa fue preparado por Nicholas Z. Scoville, F. Peter Schloerb y Paul F. Goldsmith, de la Universidad de Massachusetts en Amherst. Obtenido con un radiotelescopio de 14 metros, el mapa tiene una resolución de unos 45 segundos de arco, unas 10 veces mayor que el de la página anterior. El rectángulo delimita las regiones más luminosas de OMC1, que se muestran en mayor detalle en la página siguiente. (Cortesía de los observatorios Kitt Peak y Lick.)

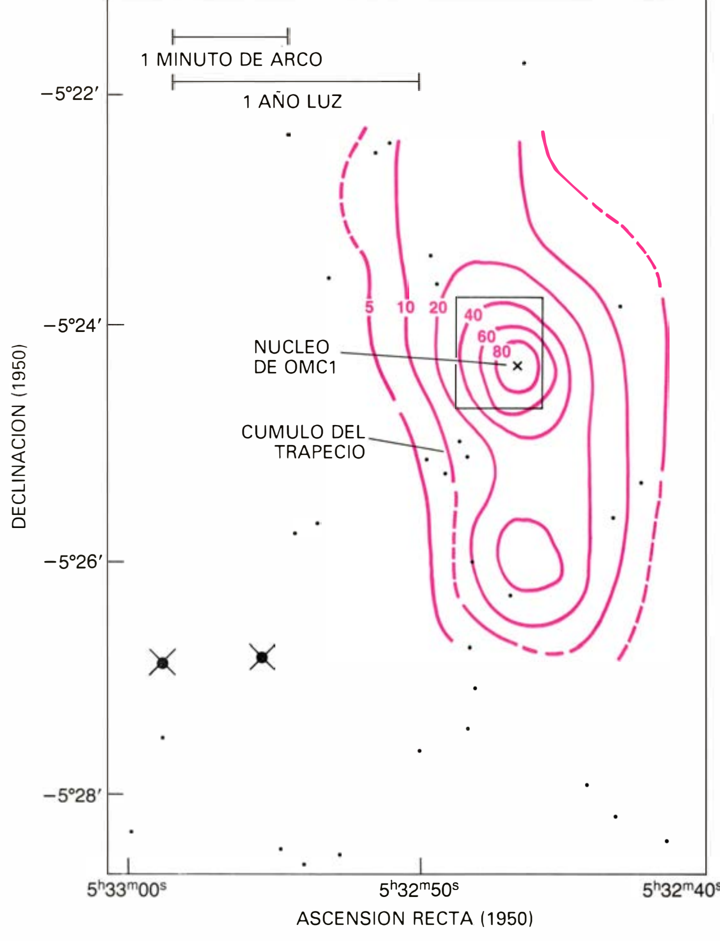
ros de ondas cuando la molécula de hidrógeno en el estado $v = 1$, $J = 2$ emite un fotón y desciende al estado fundamental $v = 0$, $J = 0$ es de 4498 centímetros inversos (cm^{-1}). El recíproco de este valor (la unidad dividida por 4498 cm^{-1}) corresponde a una longitud de onda infrarroja de 2,22 micrometros.

Si una sola molécula de una nube se dejara imperturbada, descendería rápidamente a su nivel de energía más bajo, emitiendo al hacerlo una serie de fotones. En una nube molecular real, sin embargo, las moléculas están en continua colisión unas con otras. En cada choque hay cierta probabilidad de que una molécula absorba parte de la energía cinética del impacto y sea excitada hasta uno de sus niveles energéticos superiores. Habiendo sido excitada de esta manera, emitirá entonces uno o más fotones al retornar al estado fundamental, a menos que otro choque vuelva a alterar su nivel energético. El número relativo de moléculas en los diferentes niveles depende, pues, estrictamente de la energía cinética media de las moléculas presentes y de la probabilidad de los choques. En consecuencia, las intensidades de las líneas de emisión molecular proporcionan información sobre las temperaturas y las densidades de las nubes interestelares. Mas, por desgracia, la interpretación de tales datos no suele ser ni sencilla ni unívoca.

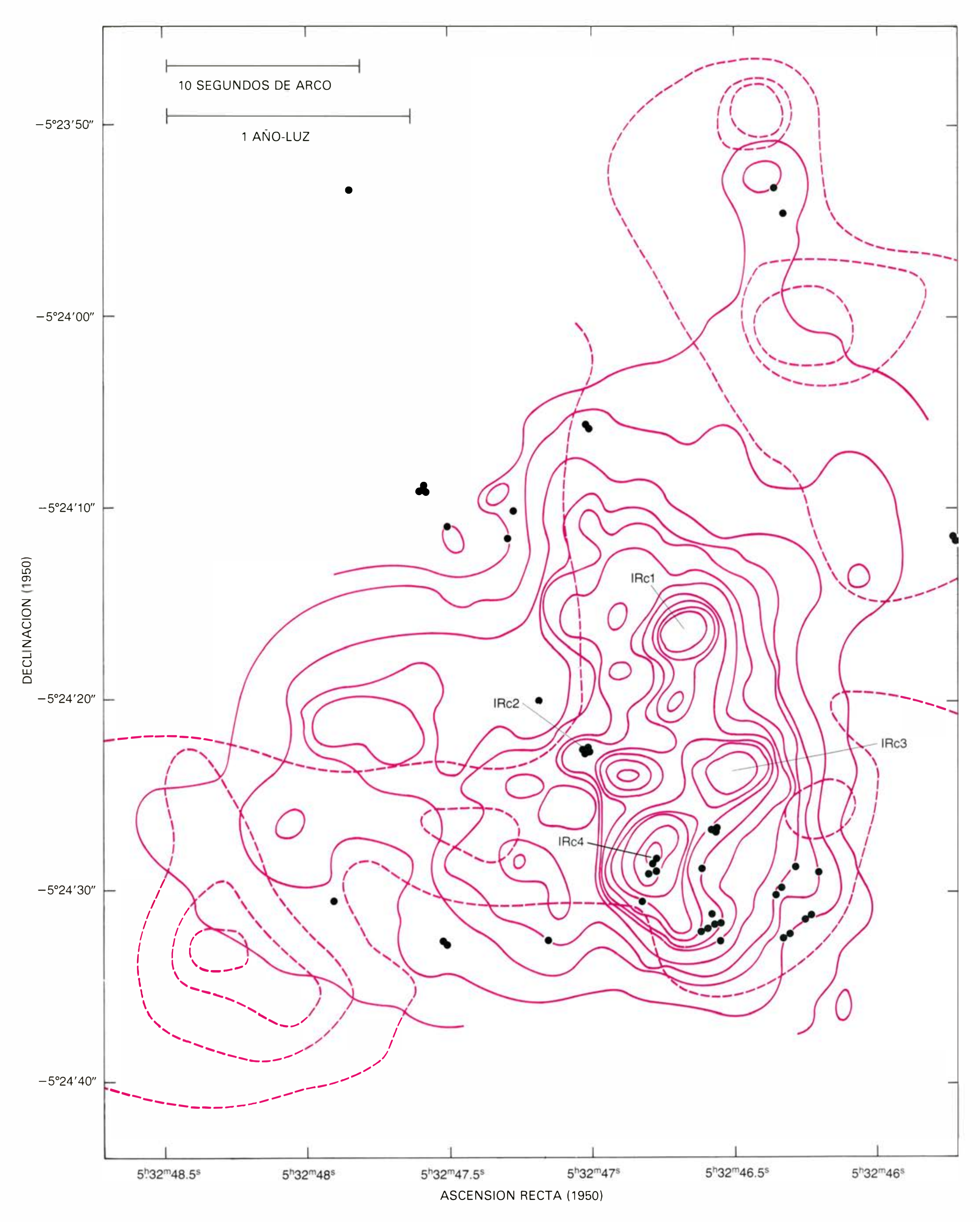
Para que las líneas espectrales del hidrógeno molecular se observaran, sería necesario, como mínimo, que una fracción apreciable de las moléculas fuera excitada hasta el estado $v = 0$, $J = 2$. (Las transiciones con $\Delta J = 1$, tales como la $v = 0$, $J = 1 \rightarrow 0$, son extremadamente raras en una molécula simétrica como la del H_2.) En las nubes interestelares típicas, que están a una temperatura de sólo unos 20 grados Kelvin, la energía cinética media de traslación (la energía del movimiento) corresponde a un fotón con un número de ondas de sólo 21 cm^{-1}. Como resultado, el estado $v = 0$, $J = 2$, que está a 354 cm^{-1} sobre el estado fundamental, se halla casi despoblado, ya que los choques raramente excitan moléculas a niveles con energías mayores que tres a cinco veces su energía cinética media. Por tanto, en las nubes interestelares típicas, esencialmente todas las moléculas de hidrógeno están confinadas a los dos niveles más bajos y, con excepción de algunas regiones calientes especiales (a las que habré de referirme), el hidrógeno molecular no puede observarse directamente en nubes como las de Orión.

Las dificultades que hacen invisible la fría molécula de hidrógeno no se aplican a la molécula de monóxido de carbono, más pesada y asimétrica. En esta molécula caben transiciones entre estados de rotación de niveles adyacentes cuya distancia energética es muy inferior a los intervalos que median entre los estados del hidrógeno. Las longitudes de onda de esas transiciones forman una progresión armónica, a partir

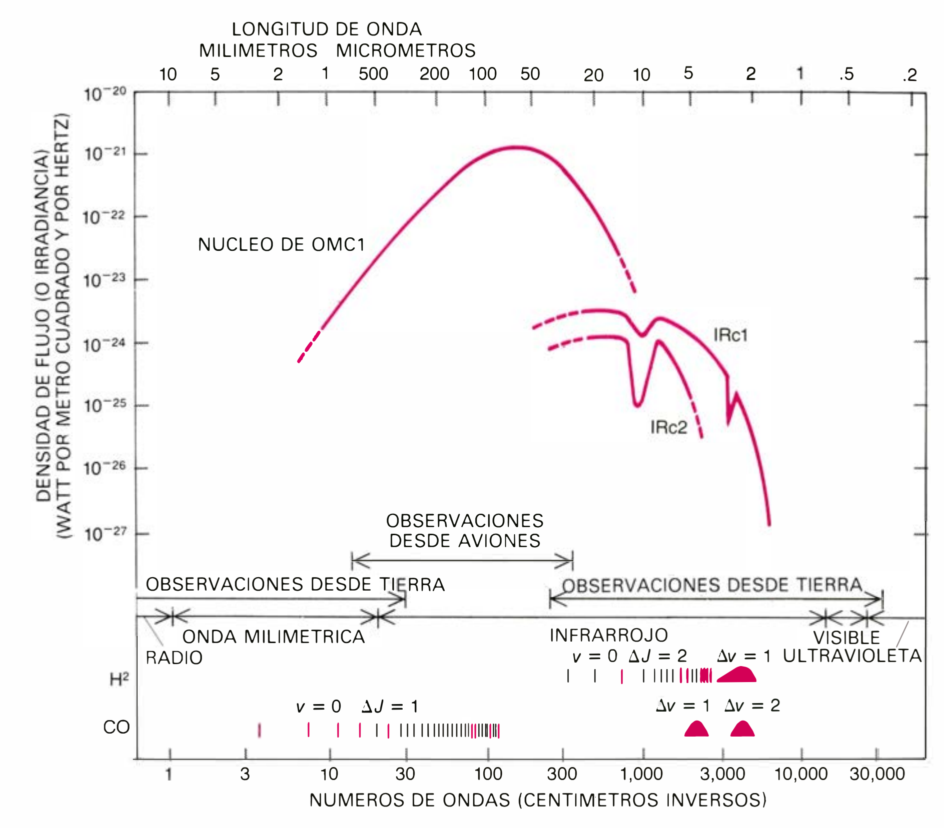




LA EMISION INFRARROJA alcanza un máximo de intensidad en la proximidad de la región visible más brillante de la Nebulosa de Orión, como puede verse comparando la fotografía en color (*izquierda*) con el mapa infrarrojo de la misma región (*derecha*). La fotografía en color fue tomada con el telescopio de tres metros del Observatorio Lick. El mapa, que muestra las isolíneas de emisión infrarroja debida a los granos de polvo calientes a la longitud de onda de 400 micrometros, se obtuvo con el equipo del telescopio infrarrojo de Mauna Kea en Hawai por Jocelyn B. Keene, Roger H. Hildebrand y Stanley E. Whitcomb, de la Universidad de Chicago. El grupúsculo de cuatro estrellas cerca del centro de la fotografía es el Cúmulo del Trapecio, agrupación de estrellas de tipo *0* cuya radiación ultravioleta ioniza el gas de la nebulosa visible y le hace emitir luz. Los granos de polvo caliente están inmersos en OMC1. Las cifras de las isolíneas representan el porcentaje de la máxima irradiancia a 400 micrometros. Los granos de polvo más calientes se encuentran en el núcleo de OMC1, lo que revela una potente fuente de energía. El rectángulo define el área que corresponde a la ilustración siguiente.



EL POLVO CALIENTE Y EL HIDROGENO MOLECULAR se concentran en una región de sólo algunas décimas de años-luz de diámetro en el núcleo de OMC1. Las curvas de trazo continuo en color muestran la distribución de polvo caliente tal como se observó a la longitud de onda de 20 micrometros con el telescopio infrarrojo de tres metros en Mauna Kea por Reinhard Genzel, Dennis Downes, Eric E. Becklin y el autor. La emisión infrarroja más intensa brota de una media docena de fuentes compactas. Entre ellas IRc1 e IRc2, cuyas temperaturas rondan los 500 grados Kelvin. Las líneas de trazos representan la emisión del hidrógeno molecular a 2,12 micrometros, una transición entre niveles de energía asociada con temperaturas del gas de alrededor de 2000 grados Kelvin. El hidrógeno se calienta probablemente por el choque entre el gas expulsado hacia el exterior de IRc2 y el gas estacionario en la nube molecular. El mapa de la emisión de hidrógeno fue realizado con el telescopio de cinco metros del Monte Palomar por Becklin, Steven Beckwith, G. Neugebauer y S. E. Persson. Becklin y Neugebauer descubrieron IRc1, u objeto BN, en 1956. Los puntos negros son los máximos de emisión de masers.



COMPARACION ENTRE LONGITUDES DE ONDA Y DISTRIBUCIONES DE ENERGIA para OMC1 y las fuentes compactas de infrarrojo designadas IRc1 e IRc2. Si se pretende cubrir la totalidad del intervalo espectral de esos objetos habrá que observar desde aviones o globos, además de centros con sede en tierra. Los números de ondas para transiciones importantes entre niveles de energía del hidrógeno molecular (H_2) y el monóxido de carbono (CO) están representados a lo largo de la parte inferior del gráfico. Algunas transiciones están tan juntas que forman bandas. Las tres bandas y las transiciones indicadas por las líneas en color se han observado en Orión. El hecho de que IRc1 e IRc2 tengan máximos de emisión a longitudes de onda más cortas que las de OMC1 implica una mayor temperatura y compacidad. Se cree que los valles de las curvas de IRc1 e IRc2 en 3,1 y 9,7 micrometros se deben a la absorción de radiación infrarroja por granos de hielo y silicatos que hay en el polvo interestelar interpuesto.

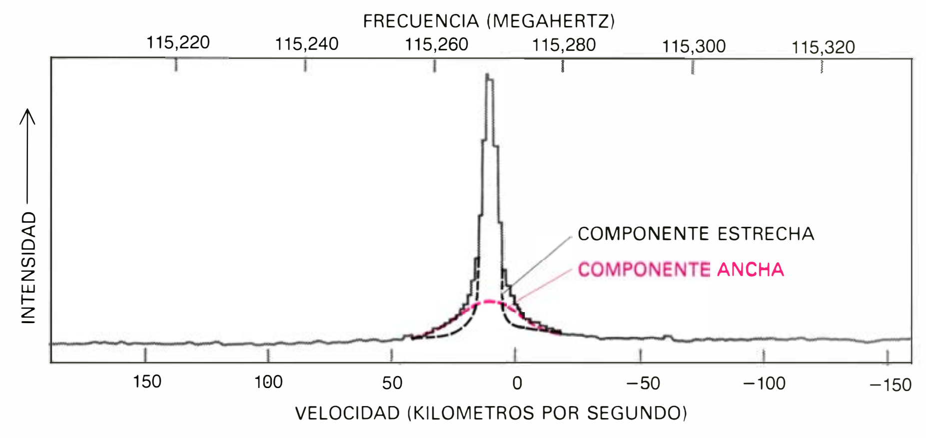
de 2,6 milímetros, con moléculas en el estado $v = 0$ que caen de $J = 1$ a $J = 0$ y prosiguen con líneas espectrales a 1,3 milímetros, 0,87 y 0,65 milímetros para transiciones más altas. Las dos transiciones de longitud de onda más larga pueden observarse con radiotelescopios de alta calidad. Las transiciones de longitudes de onda más cortas suelen requerir las más precisas superficies reflectoras de los telescopios ópticos o de infrarrojo, dotados de receptores especiales. Se han estudiado diversas moléculas por medio de sus transiciones en longitudes de onda milimétricas, pero la molécula de monóxido de carbono (presente en la relación de una molécula por cada 2000 de hidrógeno) predomina, con mucho, entre las moléculas detectadas.

Hace algunos años se realizó un estudio muy completo de la región de Orión por un grupo formado por Marc L. Kutner, del Instituto Politécnico Rensselaer, Kenneth D. Tucker, de la Universidad Fordham, Gordon Chin, de la Universidad de Columbia, y Patrick Thaddeus, del Instituto Goddard de Estudios Espaciales. Se confeccionó un mapa de toda la región, realizado por el radiotelescopio de 1,2 metros de Columbia sintonizado para detectar la más baja transición rotacional del monóxido de carbono a 2,6 milímetros. Reveló dos complejos muy grandes de nubes moleculares, cada uno de unos 150 años-luz de diámetro [*véase la ilustración de la página 23*]. Con excepción de unas pocas manchas que están iluminadas por brillantes estrellas cerca de los bordes de las nubes, la materia aparece completamente opaca en las fotografías ópticas. Aunque las nubes son densas para lo que resulta habitual en el espacio interestelar, constan sólo de algunos cientos de moléculas por centímetro cúbico. Pero la magnitud de las nubes explica que cada una contenga materia suficiente para formar 100.000 estrellas de la masa del Sol. Las mediciones del corrimiento Doppler en las líneas de emisión de las nubes indican que éstas no pueden haber alterado su configuración, de un modo relevante, en el último millón de años.

En la mayor parte del área cartografiada, las nubes están frías, por debajo de los 20 grados Kelvin; les llega únicamente el calor de la luz difusa que se filtra procedente de las estrellas vecinas. Pero las regiones que muestran la emisión de monóxido de carbono más intensa están apreciablemente más calientes: entre 50 y 100 grados. Regiones que, además, suelen poseer una densidad mayor que el resto de la nube. En el complejo situado más al Sur, el mejor estudiado de los dos, el máximo de emisión de monóxido de carbono viene a coincidir con la nebulosa de Orión [*véase la ilustración de la página 25*]. Aunque la Nebulosa de Orión y el máximo del monóxido de carbono parecen estar superpuestos, el gas brillante en la nebulosa visible está formado por átomos completamente ionizados a temperaturas de unos 10.000 grados Kelvin; la emisión de monóxido de carbono procede de la nube de hidrógeno molecular situada inmediatamente detrás de la nebulosa. A la nube no ionizada se la conoce por OMC1 (Orion Molecular Cloud 1, Nube Molecular de Orión 1). La nebulosa invade lentamente OMC1 a medida que el hidrógeno, monóxido de carbono y otras moléculas se van degradando por la radiación ultravioleta de las estrellas calientes de tipo *0* situadas dentro de la nebulosa.

Aunque parte del calentamiento de OMC1 puede atribuirse a la energía absorbida de las estrellas de tipo *0*, el máximo más intenso en la emisión del monóxido de carbono requiere otra explicación. Examinando las líneas de emisión de diversas clases de moléculas, cada una de ellas afectada de una manera diferente por la densidad y temperatura, los astrónomos han hallado que OMC1 tiene un núcleo de diámetro pequeño, de algunas decenas de años-luz, donde la densidad supera las 100.000 moléculas por centímetro cúbico. La masa de gas en el núcleo se estima entre algunos cientos y algunos miles de veces la masa solar. La fuerte concentración de gas en este punto, la indicación de que debe haber algo dentro de la nube que la calienta y el marcado ensanchamiento de las líneas de emisión (prueba de que el movimiento es allí más activo que en el resto de la nube), todo ello sugiere que el núcleo es la sede de formación muy reciente de estrellas o bien de estrellas en ciernes.

La observación de transiciones moleculares de longitudes de onda milimétrica en nubes interestelares presenta dos limitaciones. En primer lugar, tales transiciones no revelan nada acerca de



VELOCIDADES DEL GAS EN OMC1, inferibles a partir de la magnitud del corrimiento Doppler en la línea de emisión del monóxido de carbono a 2,6 milímetros, equivalente a una frecuencia de 115.270 megahertz. En la mayoría de los lugares de la nube gaseosa, la emisión está confinada a una estrecha línea (*negra*), pero en el núcleo se observa una componente adicional con un perfil ancho (*color*). El ensanchamiento indica que las velocidades del gas están comprendidas entre −50 y +50 kilómetros por segundo. La componente ancha se halla confinada a una región cercana a IRc2 de sólo 40 segundos de arco.

la naturaleza de las estrellas o de las otras fuentes de energía que puedan estar escondidas dentro de la nube. En segundo lugar, el método adolece de una pobre resolución angular; los detalles más pequeños que pueden cartografiarse mediante un telescopio dependen de la relación entre la longitud de onda de la radiación y el diámetro del telescopio. La resolución que se puede alcanzar, cuando la longitud de onda de 2,6 milímetros del monóxido de carbono se recoge en un telescopio de 14 metros de radio, es de 45 segundos de arco, es decir, unas 20 veces menos de lo que puede conseguirse con telescopios en longitudes de onda visibles. Las observaciones en las longitudes de onda infrarrojas entre 0,7 micrometros y 500 micrometros (0,5 milímetros) están ayudando a salvar ambas restricciones.

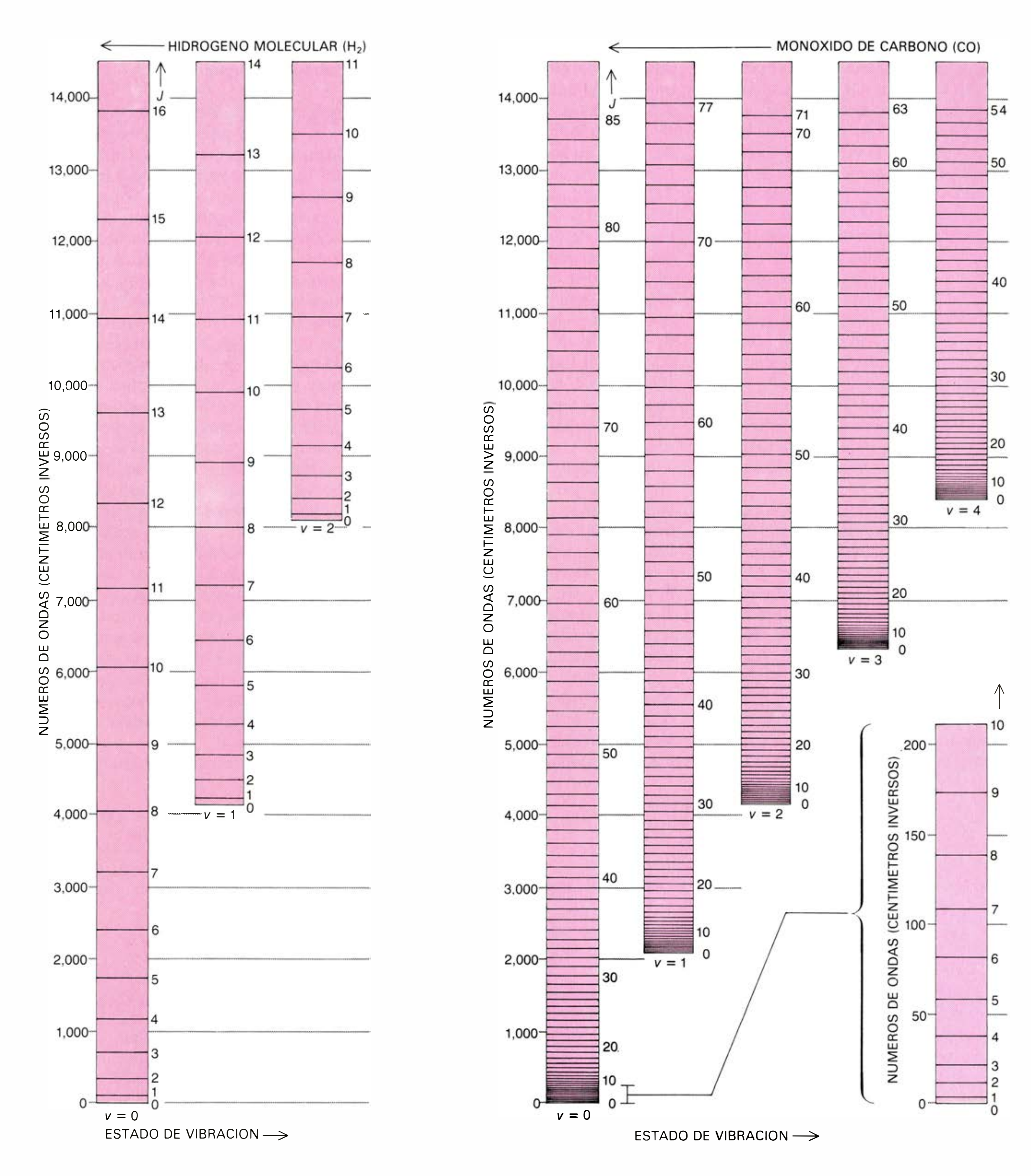
La principal fuente de radiación infrarroja cósmica es la radiación térmica procedente de granos de polvo interestelar calientes. Aunque la composición y tamaño de éstos sea incierta, incluyen probablemente silicatos y gases congelados, y, al menos, algunos granos deben de tener diámetros de alrededor de 0,1 micrometros. Los granos poseen una elevada eficacia para absorber la luz de las estrellas y reemitir su energía en forma de radiación infrarroja. La longitud de onda de la radiación emitida depende de la temperatura de los granos. A 500 grados Kelvin, por ejemplo, la mayor parte de la radiación queda entre las longitudes de onda de cinco y 15 micrometros. A temperaturas más bajas dominarían mayores longitudes de onda, siguiendo la bien conocida ley de Planck para la radiación del cuerpo negro.

En la mayoría de las nubes interestelares, los granos de polvo están bien distribuidos y contribuyen, en un 0,5 por ciento aproximadamente, al valor total de la masa de la nube. Jocelyn B. Keene, Roger H. Hildebrand y Stanley E. Whitcomb, de la Universidad de Chicago, cartografiaron la radiación procedente de granos calientes en OMC1, a la longitud de onda de 400 micrometros, con el Telescopio Infrarojo de tres metros de la Universidad de Hawai, instalado a una altura de 4150 metros en Mauna Kea. A esta longitud de onda bastante larga y baja resolución (unos 35 segundos de arco), la radiación infrarroja parece originarse en una región toscamente elíptica, de unos dos años-luz en su eje mayor, que comprende el núcleo de OMC1. Cuando la fuente infrarroja se observa a longitudes de onda más cortas, en las que domina la radiación de los granos más calientes, se encuentra que la radiación se concentra, casi por entero, en el interior del núcleo de OMC1, lo que parece indicar que éste aloja la fuente de energía que calienta la nube. Medidas realizadas desde un avión en vuelo, a una altura superior a 12 kilómetros (para reducir la absorción de radiación infrarroja en la atmósfera terrestre), muestran que la mayor parte de la radiación procedente del núcleo de OMC1, que importa unas 100.000 veces la potencia radiada por el Sol, se emite entre 20 y 300 micrometros (intervalo de longitudes de onda para el cual la atmósfera terrestre es casi completamente opaca).

Se acaba de obtener un mapa con resolución mucho más alta que el de 400 micrometros (dos segundos de arco comparados con 35 segundos de arco) por un grupo integrado por Reinhard Genzel, del Centro de Astrofísica del Observatorio del Harvard College y del Observatorio Astrofísico Smithsoniano, Dennis Downes, del Instituto alemán Max Planck de Radioastronomía, Eric E. Becklin, de la Universidad de Hawai, y el autor con el sistema telescópico infrarrojo de tres metros en Mauna Kea, a una longitud de onda de 20 micrometros [*véase la ilustración de la página 26*]. El mapa revela que la emisión infrarroja procede de un cúmulo de media docena de objetos, o más, rodeados por emisión difusa. A partir de la variación de la irradiancia infrarroja con la longitud de onda podemos estimar la temperatura aparente de los objetos. Los más calientes, designados IRc1 e IRc2 (fuente compacta infrarroja Nos. 1 y 2), radian intensamente en el intervalo comprendido entre 2 y 20 micrometros, lo que indica una temperatura de unos 500 grados Kelvin. Otros varios máximos, que incluyen IRc3 e IRc4, se hallan sólo a unos 150 grados. Todos estos objetos tienen temperaturas muy por debajo de la temperatura de la más fría de las estrellas normales; por tanto, deben representar concentraciones de polvo caliente. Resulta difícil calcular la potencia total radiada por cada concentración, en razón de los efectos atenuadores del polvo situado entre emisor y observador, pero cabe presumir que supere en 10.000 veces la emisión solar.

¿Cuál es la naturaleza de las fuentes compactas de infrarrojo situadas dentro de OMC1? Poco después de su descubrimiento se supuso que podría tratarse de protoestrellas, es decir, objetos intermedios, en el curso de la evolución, entre nubes interestelares y estrellas propiamente dichas. Los primeros cálculos detallados sobre la evolución y aparición de protoestrellas se deben a Richard B. Larson, de la Universidad de Yale, en 1969. Su modelo teórico muestra que una nube, contrayéndose bajo la acción de su propia gravitación, forma, inicialmente, un núcleo de materia de alta densidad en su centro. Al caer más materia en el centro, la energía gravitatoria liberada calienta los granos de polvo de la capa exterior de la protoestrella, haciéndoles emitir radiación infrarroja. Mientras tanto, la temperatura y densidad del núcleo aumentan de continuo hasta llegar a desencadenar reacciones termonucleares en el centro. Para una protoestrella con la masa del Sol, el proceso entero dura al menos 10 millones de años; para protoestrellas con la masa de las estrellas de tipo *O* y de tipo *B*, sin embargo, el proceso resulta hasta 100 veces más rápido.

Trabajos posteriores revelaron que,



TRANSICIONES ENTRE NIVELES DE ENERGIA de las moléculas diatómicas de hidrógeno y monóxido de carbono. Dichas transiciones dan lugar a la radiación que ha suministrado gran parte de la información acerca de las nubes gaseosas en la vecindad de la Nebulosa de Orión. Las moléculas diatómicas pueden existir en muchos niveles de energía por encima del nivel fundamental, que es el estado de vibración y rotación de energía mínima para una molécula. En ese estado, tanto el número cuántico de vibración, v, como el número cuántico de rotación, J, son ambos nulos. En una nube molecular, el H_2 y el CO pueden adquirir energía y ser excitados, por choques con otras moléculas, a estados de más alta energía de vibración y de rotación. Para cada estado de vibración hay muchos estados de rotación posibles. En general, se necesita más energía para efectuar una transición entre estados de vibración con el mismo valor de J que la que hace falta para efectuar una transición entre estados de rotación con el mismo valor de v. Cuando una molécula pasa de un estado de energía a otro inferior libera un cuanto de radiación en forma de fotón. La longitud de onda del fotón emitido, expresada en centímetros, es el recíproco de la diferencia entre los niveles energéticos en número de ondas, representada a lo largo del eje vertical. Por ejemplo, una molécula de hidrógeno en el estado $v = 1$, $J = 2$ tiene un número de ondas de 4498 centímetros inversos. Si la molécula descendiera al estado fundamental, $v = 0$, $J = 0$, en el que el número de ondas es cero, liberaría un fotón con una energía de 1/4498 centímetros o 2,22 micrometros. En algunas de las regiones más calientes de OMC1, las moléculas de hidrógeno emiten radiación a esa longitud de onda. Para el H_2, cuya estructura molecular es simétrica, las transiciones entre niveles adyacentes de rotación están prohibidas. Por tanto, en nubes frías de gas no hay energía suficiente para elevar las moléculas de hidrógeno a niveles de energía en los que se puedan observar transiciones. Esto no se aplica a las moléculas de CO, que son asimétricas y libres para moverse entre niveles adyacentes de rotación. En las nubes moleculares frías, se pueden observar muchas transiciones de moléculas de CO en el estado $v = 0$ con energías rotacionales entre $J = 0$ y $J = 6$. La transición más baja del CO, de $J = 1$ a $J = 0$, da un fotón con una longitud de onda de 2,6 milímetros.

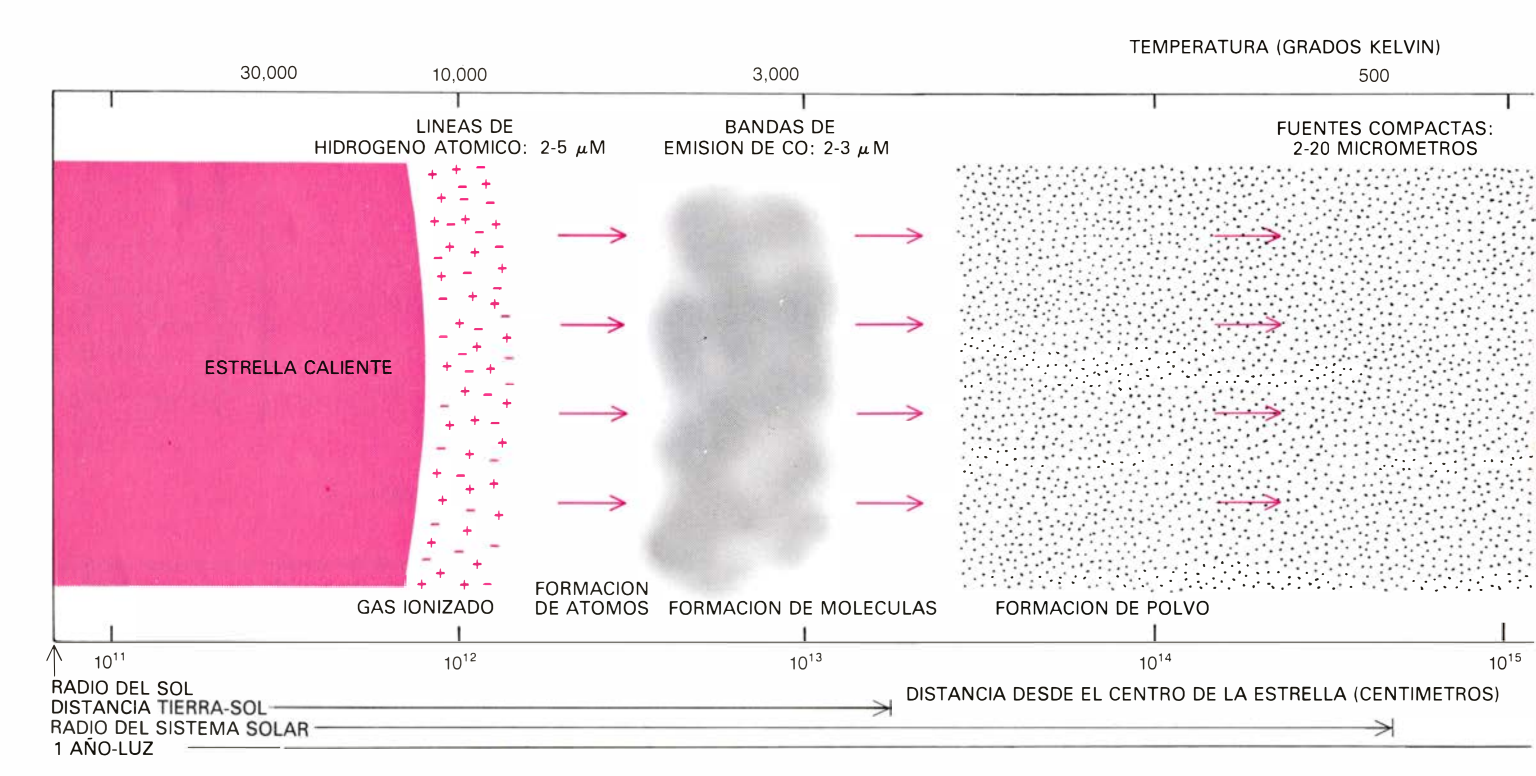
en las protoestrellas de gran masa, las reacciones termonucleares comenzaban antes de que las regiones externas tuvieran tiempo de caer en el núcleo. La estrella recién nacida estaría rodeada por un capullo de gas y polvo, emisor de infrarrojo, en proceso de colapso. Si las fuentes infrarrojas compactas observadas en OMC1 eran precursoras de estrellas de tipo *0* y *B*, quedaría explicada su luminosidad. De acuerdo con este esquema, el cúmulo infrarrojo acabaría con el tiempo por aparecer como un pequeño grupo de estrellas *0* y *B*, de tamaño y luminosidad semejantes a las de las estrellas del Cúmulo del Trapecio situadas dentro de la propia Nebulosa de Orión. El resto de la nube molecular terminaría por calentarse hasta miles de grados y comenzaría a brillar como una nueva nebulosa ionizada.

Pero las objeciones a ese razonamiento se presentaron ya en 1976, al descubrirse, en el núcleo de OMC1, gas que se movía a gran velocidad. El mérito del hallazgo hay que atribuírselo a Ben M. Zuckerman, de la Universidad de Maryland en College Park, Thomas B. H. Kuiper, del Instituto de Tecnología de California, Eva N. Rodríguez Kuiper, de la Universidad de California en Los Angeles, John Kwan, de la Universidad estatal de Nueva York en Stony Brook, y Nicholas Z. Scoville, de la Universidad de Massachusetts. Observaron corrimientos Doppler, en la emisión del monóxido de carbono procedente de la sede del cúmulo infrarrojo, que indican velocidades del gas de más o menos 50 kilómetros por segundo con respecto al resto de la nube. Tales velocidades, más de 10 veces más altas que las observadas en cualquier otra parte de OMC1, superan con mucho la celeridad que podría inducir el mero colapso gravitatorio. Los movimientos a gran velocidad están confinados en un pequeño volumen, de unos 0,3 años-luz de diámetro, y se centran dentro de 0,1 años-luz de IRc2. La distribución de las velocidades Doppler sugiere una capa gaseosa que se expande en todas las direcciones a partir de una fuente.

En ese mismo año de 1976 se aportó una nueva prueba a favor del movimiento rápido en dicha región: el descubrimiento de líneas de emisión en el infrarrojo procedentes de moléculas calientes de hidrógeno que realizaron T. N. Gautier III, Uwe Fink, Richard R. Treffers y Harold P. Larson, de la Universidad de Arizona. Trabajaban con el telescopio de 2,29 metros del Observatorio Steward. Desde entonces se han observado más de una docena de transiciones en el hidrógeno molecular, la mayoría de ellas entre niveles bajos de *J* de los estados de vibración $v = 1$ y $v = 0$ a longitudes de onda próximas a los dos micrometros.

Hay cuatro aspectos a destacar en torno a esas emisiones. En primer lugar son intensas, con una potencia total que centuplica, cuando menos, la potencia del Sol a todas las longitudes de onda. En segundo lugar, y tal como se deduce de la relación de intensidades entre distintas líneas de emisión, la temperatura del hidrógeno alcanza o supera los 2000 grados Kelvin, mucho más alta que el resto del gas en OMC1. En tercer lugar, los ensanchamientos Doppler de las líneas espectrales son muy grandes, lo que implica velocidades entre 100 y 150 kilómetros por segundo. En cuarto lugar, el gas caliente muestra una amplia distribución en la región de OMC1, tomando la forma de bolsas alrededor del cúmulo infrarrojo. Todas esas observaciones sugieren que la emisión del hidrógeno molecular resulta de un "viento" de gran intensidad soplando hacia el exterior del cúmulo infrarrojo. Cuando este gas, que probablemente está compuesto de la misma materia que da lugar a las anchas emisiones de monóxido de carbono, choca con las partes circundantes de OMC1, crea una onda de choque que calienta, hasta 2000 grados, una delgada y densa capa de OMC1.

Una tercera clase de pruebas, que

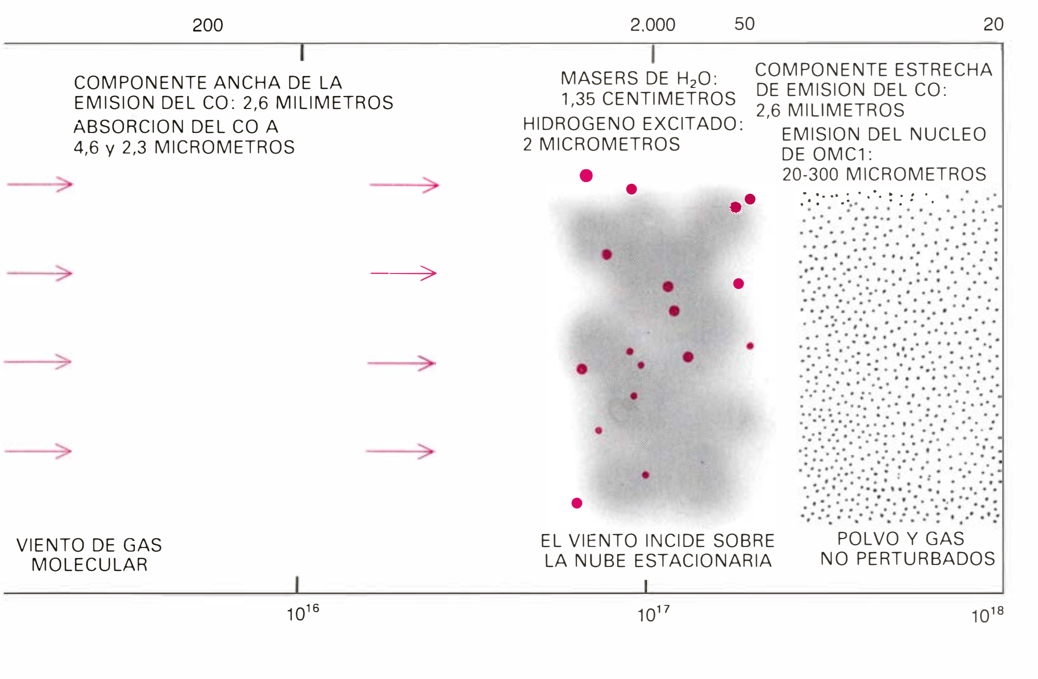


CONDICIONES HIPOTETICAS en la vecindad de una estrella recién formada en OMC1, deducidas a partir de observaciones de IRc1 e IRc2. Las observaciones son casi todas explicables sólo si se supone que un potente "viento" está soplando desde una estrella que está oculta en el centro de cada fuente compacta. Presumiblemente, en la nueva estrella se han iniciado las reacciones termonucleares, calentando su superficie hasta unos 3000 grados Kelvin. La radiación ultravioleta de la estrella ioniza el gas, principalmente hidrógeno, en la capa circundante. La capa absorbe toda la radiación ultravioleta y produce líneas de emisión de una pequeña fracción de átomos de hidrógeno que están calientes pero no ionizados. A medida que el gas expulsado se enfría, los átomos se combinan formando moléculas, entre ellas las de hidrógeno y monóxido de carbono. El último emite radiación a 2,3 micrometros. El

apoyan la existencia de movimientos expulsivos en Orión, proviene de la observación de intensas emisiones de masers de agua. La palabra maser significa amplificación de microondas por emisión estimulada de radiación. Un maser funciona elevando un gran número de átomos o moléculas de un gas a un estado particular de energía y entonces estimulándolos a caer simultáneamente a algún estado más bajo. Como la transición al estado más bajo se produce simultáneamente, la cantidad de fotones emitidos por unidad de tiempo por un volumen dado de un gas de una densidad dada es enormemente mayor que la correspondiente cantidad debida a los procesos cinéticos aleatorios que he descrito hasta ahora. En un maser de agua, el gas está constituido por moléculas de agua presentes, en pequeñas cantidades, en el seno del gas hidrógeno. La primera emisión de un maser de agua celeste se observó a la longitud de onda de 1,35 centímetros en 1969. Desde entonces, se han descubierto otras en más de 100 regiones de la galaxia. En OMC1 hay al menos 50 nubecillas de maser, distribuidas sobre regiones de aproximadamente la misma extensión que el cúmulo infrarrojo. La emisión de un maser de agua celeste a 1,35 centímetros puede ser extremadamente intensa. Para emitir una cantidad igual de energía, a esa longitud de onda, mediante un proceso térmico ordinario, el cuerpo tendría que estar a una temperatura de más de 10^{12} grados Kelvin.

La alta intensidad de la emisión de la línea de 1,35 centímetros y la compacidad de las fuentes de masers de agua posibilitan el que los radioastrónomos puedan medir, con gran precisión, las distancias entre las diversas nubecillas de maser; para ello recurren a la interferometría de muy larga base. Se realizaron observaciones simultáneas de la región de Orión con radiotelescopios en California, Viginia Occidental, Massachusetts, Alemania, Suecia y la Unión Soviética por un grupo en el que figuraban Downes, Genzel, Mark J. Reid y James M. Moran, los dos últimos del Centro de Astrofísica. Observando a lo largo de un período de dos años, el grupo pudo seguir pequeñas variaciones en las distancias que separan las nubecillas de maser en Orión. Con la interferometría de muy larga base se pudieron medir los ángulos de separación con una precisión de 10^{-3} segundos de arco, que viene a representar el ángulo subtendido por una moneda pequeña a una distancia de 4000 kilómetros.

Las mediciones revelaron que muchas de las nubecillas de maser de agua se movían, alejándose de un punto que está en algún lugar en la proximidad de IRc2 o IRc4, con una velocidad de 18 kilómetros por segundo. La identificación de IRc2 como la fuente de la corriente hacia fuera vino sugerida por el descubrimiento de un segundo tipo de emisión maser, de moléculas de monóxido de silicio, coincidiendo estrechamente con IRc2 y presentando una velocidad de expansión semejante a la de los masers de agua. Algunas de las regiones de maser de agua se están alejando de IRc2 a velocidades de hasta 100 kilómetros por segundo, pero no podemos afirmar todavía que formen parte de la misma corriente que propulsa los masers de agua y de monóxido de silicio, más lentos, o que emanen de una de las otras fuentes infrarrojas.



polvo, calentado directamente por la estrella, constituye la fuente de radiación infrarroja a longitudes de onda entre 2 y 20 micrometros. Entre 10^{16} y 10^{17} centímetros del centro de la estrella, el monóxido de carbono, que se mueve velozmente con el viento, emite radiación a una longitud de onda de algunos milímetros. A unos 10^{17} cm, donde el viento incide sobre la nube circundante, las moléculas de hidrógeno se calientan nuevamente hasta 2000 grados K, dando lugar a emisiones a unos dos micrometros. Quizá como parte del fenómeno de la colisión, las nubecillas forman masers de agua que radian a 1,35 cm.

En cualquier caso, la cantidad de materia que participa en la corriente desde IRc2 es importante; equivale a una pérdida de entre una milésima y diez milésimas partes de la masa del Sol cada año por un objeto que no es probable supere las 20 masas solares de materia. Hasta ahora se desconoce la causa de esa enorme velocidad de pérdida de materia, pérdida que es unas 100 veces mayor que la de las estrellas maduras de tipo *0* y no predicha por ninguna de las actuales teorías de formación de estrellas. Entre las posibilidades sometidas a investigación se cuentan la pérdida de masa por presión de radiación, por rotación, por ondas acústicas, por efectos magnéticos y por turbulencia inducida por el viento en el disco circumestelar de una estrella recién formada (y quizás incluso un disco preplanetario).

Cualquiera que sea el mecanismo responsable de la pérdida de masa, resulta claro que una fracción notable de la actividad observada en el núcleo de OMC1 puede atribuirse directamente a procesos que ocurren dentro de una o varias fuentes compactas de infrarrojo. ¿Qué esperanzas hay de examinar tales procesos en mayor detalle y de descubrir la naturaleza y fase de evolución de dichos objetos? Una potente herramienta, de desarrollo reciente, es el espectrómetro de transformada de Fourier. Este instrumento consigue espectros detallados de fuentes infrarrojas individuales, a longitudes de onda entre uno y cinco micrometros, con una resolución que corresponde a una velocidad Doppler de algunos kilómetros por segundo. Desgraciadamente, IRc2 es débil en tal intervalo de longitudes de onda, pero se han realizado estudios de la vecina fuente infrarroja IRc1, muy brillante, con el telescopio de cuatro metros de Kitt Peak por Donald N. B. Hall, Stephen T. Ridgway y Fred C. Gillett, del Obser-

vatorio Nacional de Kitt Peak, asociados con Susan G. Kleinmann, del Instituto de Tecnología de Massachusetts, y Scoville.

Su espectro de IRc1 muestra líneas de emisión y de absorción. Las moléculas de monóxido de carbono, en varios estados de rotación del estado fundamental de vibración ($v = 0$), aparecen excitadas hasta niveles en los estados $v = 1$ y $v = 2$, dando lugar a dos series de líneas de absorción muy juntas, una cerca de 4,6 micrometros y la otra cerca de 2,3 micrometros. La absorción era de esperar; se debe al gas de la parte de OMC1 que queda delante de IRc1. Menos esperado fue el descubrimiento de que la existencia de gas adicional absorbente alejándose, desde IRc1 hacia nosotros, a unos 40 kilómetros por segundo, presumiblemente parte de un viento que procede de la fuente infrarroja o una vecina próxima.

En la banda de dos a cinco micrometros del espectro de IRc1 se observan también muchas líneas de emisión. Algunas de ellas representan transiciones entre niveles de energía de átomos —no de moléculas— de hidrógeno, y son idénticas a las líneas emitidas por nubes interestelares que, como la propia nebulosa de Orión, se mantienen en un estado casi completamente ionizado en virtud de la radiación ultravioleta de las estrellas de tipo *0* y de tipo *B*. El hecho de que IRc1, a diferencia de la Nebulosa de Orión, no emita continuo de radiación (esto es, un fondo de radiación distinto de las líneas) a longitudes de onda de algunos centímetros implica que la región ionizada es muy compacta, con un diámetro de menos de 10 horas-luz. El continuo de radiación se genera en un gas ionizado por la desviación de partículas cargadas al ser aceleradas o frenadas cuando se acercan a otras partículas. Al no emerger de IRc1 tal radiación, debe indicar que se absorbe antes de poder escapar.

La radiación infrarroja de IRc1 comprende asimismo muchas líneas de emisión del monóxido de carbono. De las estudiadas hasta ahora, las más intensas proceden de transiciones en las que v varía en dos unidades, es decir, las transiciones $v = 4 \rightarrow 2$, $v = 3 \rightarrow 1$ y $v = 2 \rightarrow 0$. En esas transiciones hay implicados muchos niveles J, lo que tiende a juntarlas y desdibujarlas. Esto sucede en particular cerca de $J = 50$ en cada nivel de vibración, generando, más que líneas discretas, bandas en el espectro. Lo que sorprende, en relación con las líneas de emisión, que aparecen en longitudes de onda de alrededor de 2,3 micrometros, es que deban originarse en regiones donde los niveles muy altos de energía del monóxido de carbono (al menos el nivel $v = 4$, $J = 50$) están abundantemente poblados. Esto sólo puede suceder si la temperatura supera los 3000 grados Kelvin y la densidad, las 10^{10} moléculas por centímetro cúbico (lo que es denso para un gas interestelar, pero sólo una milmillonésima parte de la densidad de la atmósfera terrestre al nivel del mar). Bajo esas condiciones, un gas debe hallarse lindante con su fuente de calor, más cerca incluso que el polvo a 500 grados que da lugar a la emisión infrarroja entre 2 y 20 micrometros.

Aunque los objetos infrarrojos compactos IRc1 e IRc2 no son, en manera alguna, idénticos en sus propiedades observadas, tienen lo suficiente en común para sugerir una interpretación de lo que está sucediendo al menos en algunas de las fuentes infrarrojas detectadas en el interior de OMC1 y en otras nubes moleculares de apariencia semejante. Tales objetos han trascendido, probablemente, la fase de deposición que justificaría el uso del término protoestrella, lo que no debilita la creencia de que las fuentes infrarrojas están asociadas a los estados más primitivos de la evolución estelar. Su existencia como grupo compacto de objetos, su asociación con densas nubes interestelares y sus altas velocidades de pérdida de masa descartan definitivamente la posibilidad de que se trate de estrellas en las fases finales de la evolución que hayan seguido un curso distinto de las de la banda principal.

Las indicaciones disponibles sugieren que, en el centro del objeto infrarrojo, debe haber una estrella caliente no muy diferente de una estrella de la banda principal de los tipos *0* o *B*. En condiciones normales, tal estrella emitiría suficiente radiación infrarroja para mantener a su alrededor un importante volumen de gas en un estado luminoso ionizado, como es el caso de la Nebulosa de Orión. En el caso del objeto infrarojo, sin embargo, la estrella o el disco de materia en órbita a su alrededor está expulsando tanta materia en forma de un denso viento, que todos los fotones ultravioleta quedan absorbidos en una capa de gas muy próxima a la estrella. La capa se ioniza en su casi totalidad por los fotones de alta energía y origina las líneas de emisión del hidrógeno atómico observadas.

Más allá del alcance de la radiación ultravioleta, el gas continúa moviéndose hacia fuera; a medida que baja su temperatura, los átomos del gas pueden combinarse entre sí para formar moléculas y granos de polvo. Las moléculas calientes de monóxido de carbono dan lugar a la banda de emisión observada alrededor de 2,3 micrometros. El polvo, que se calienta hasta temperaturas de 500 grados Kelvin por radiación directa de la estrella (ahora a algunas horas-luz de distancia, lo que equivale al radio del sistema solar), emite radiación infrarroja a longitudes de onda entre dos y 20 micrometros. Más allá de la capa de polvo, el viento, ahora enfriado hasta temperaturas muy por debajo de los 500 grados, arrastra moléculas de monóxido de carbono que producen una amplia banda de emisión a longitudes de onda de algunos milímetros. A una distancia de alrededor de un mes-luz de la estrella recién nacida, el viento termina por chocar con la parte relativamente imperturbada de la nube molecular, originando las líneas de emisión en el infrarrojo de moléculas de hidrógeno calentadas hasta alrededor de 2000 grados. Los masers de agua pueden estar inmersos en el viento mismo o tratarse de un fenómeno relacionado con la incidencia del viento sobre la nube circundante.

El descubrimiento de que al menos algunas de las fuentes infrarrojas que antes se suponían protoestrellas son, más probablemente, estrellas muy jóvenes y de gran masa que están expulsando masa de manera espectacular, tiene algunas importantes implicaciones en nuestra concepción del proceso de formación de las estrellas. En primer lugar, significa que los astrónomos tienen que comenzar de nuevo a buscar los precursores de las estrellas típicas de la banda principal. En segundo lugar, el viento procedente de una estrella grande y luminosa puede ejercer un fuerte influjo, positivo o negativo, sobre la creación de pequeñas estrellas, tales como las semejantes al Sol. Por un lado, el viento podría deformar también la nube que lo rodea, de suerte que resultara inviable la formación ulterior de estrellas. Por otro, la presión del viento sobre las partes vecinas de la nube podría provocar el colapso de más fragmentos. En tercer lugar, si un fuerte viento es una característica de la primitiva evolución de todas las estrellas, y no sólo de las estrellas de gran masa, podría condicionar negativamente la formación de sistemas planetarios. Pero no es fácil estimar el alcance del efecto real. Si algunos astrónomos confiesan haber sufrido una desilusión al advertir que la búsqueda de protoestrellas vuelve a ser poco concluyente, pueden mitigar su frustración sabiendo que tal vez han dado con una nueva fase en la evolución de estrellas de gran masa.

La enfermedad holandesa del olmo

Esta infección fúngica, de consecuencias mortales para los olmos, puede controlarse ahora con ayuda de nuevas técnicas biológicas que atacan, simultáneamente, al hongo y al escarabajo que lo propaga

Gary A. Strobel y Gerald N. Lanier

La enfermedad holandesa del olmo es uno de los males vegetales más devastadores del siglo XX. Ha matado ya millones de olmos en Europa y Norteamérica, causando pérdidas directas a la economía por valor de miles de millones de dólares y ocasionando un perjuicio inestimable sobre la calidad estética del paisaje. Sus causas se conocen desde hace muchos años: un hongo, que se transporta de un árbol a otro, principalmente por medio de escarabajos de la corteza contaminados. Sin embargo, sólo en los últimos años se ha progresado en la comprensión detallada de la epidemiología y la química de la enfermedad, hasta el punto de que hay fundadas esperanzas de reducir de forma significativa la tasa de muerte de los olmos atribuida a la infección fúngica, mediante un programa integrado de control de la plaga. Las técnicas más modernas en este esfuerzo incluyen dos líneas muy prometedoras de control biológico: consiste la primera en manipular el comportamiento del escarabajo de la corteza con la ayuda de compuestos químicos fundamentales para su comunicación y reproducción; la segunda, en inocular los árboles con una bacteria que inhibe el crecimiento del hongo.

Fue una botánica holandesa, Dina Spierenburg, la primera que llamó la atención sobre la enfermedad, en su país, en 1919. Al poco, también en Francia y Bélgica comenzó a advertirse la muerte de olmos. Algunos observadores atribuyeron la enfermedad a los gases venenosos utilizados durante la primera guerra mundial. Otros pensaron que se debía al ataque del escarabajo de la corteza, favorecido por la sequía o por bacterias patógenas.

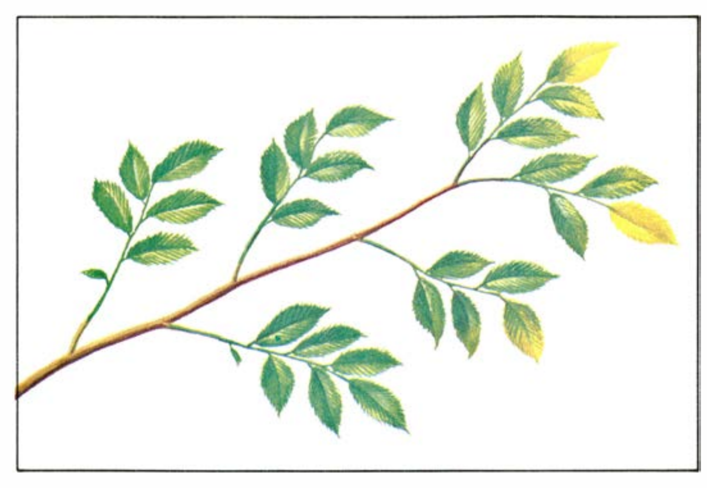
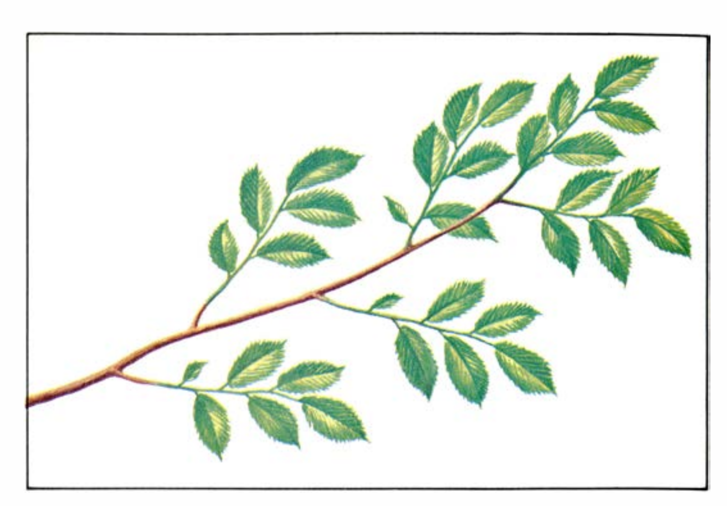
Marie Beatrice Schwarz, también de los Países Bajos, anunció en 1921 que había descubierto un hongo asociado con la madera de olmos enfermos. Cuando inyectó el supuesto patógeno en ejemplares sanos, los árboles manifestaron síntomas de la enfermedad; logró, además, volver a aislarlo de la madera que había infectado. Había identificado el agente causal mediante lo que los patólogos denominan "postulados de Koch" (en memoria del pionero de la bacteriología, el alemán Robert Koch).

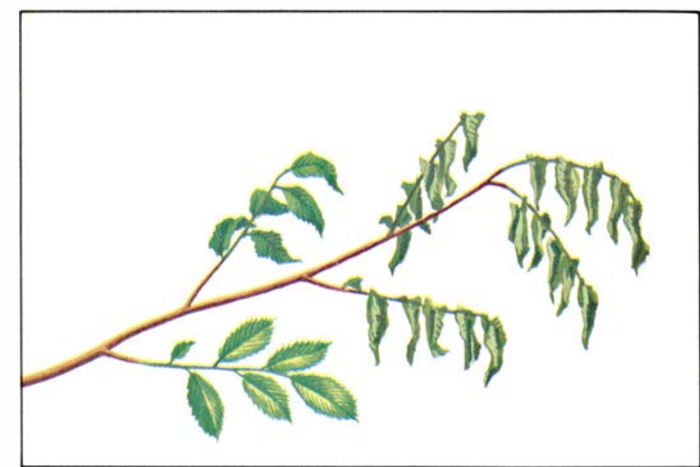
Otro punto obscuro a dilucidar en la enfermedad del olmo era la manera como penetraba el hongo en el sistema vascular del árbol. Varios investigadores europeos sospechaban que los escarabajos de la madera de la familia Scolytidae, que se encontraban habitualmente reproduciéndose en el floema, o corteza interna de olmos enfermos, diseminaban la infección. Fue, sin embargo, William Middleton, de la Oficina de Entomología y Cuarentena Vegetal del Departamento de Agricultura de los Estados Unidos, quien demostró esta suposición. En 1934 informó que la infección aparecía donde las horcaduras de las ramas resultaban heridas al alimentarse los escarabajos europeos de la corteza, de la especie *Scolytus multistriatus*, contaminados de hongos. Más tarde se supo que otras especies, europeas y americanas, podían actuar como vectores, o portadores del hongo responsable de dicho mal.

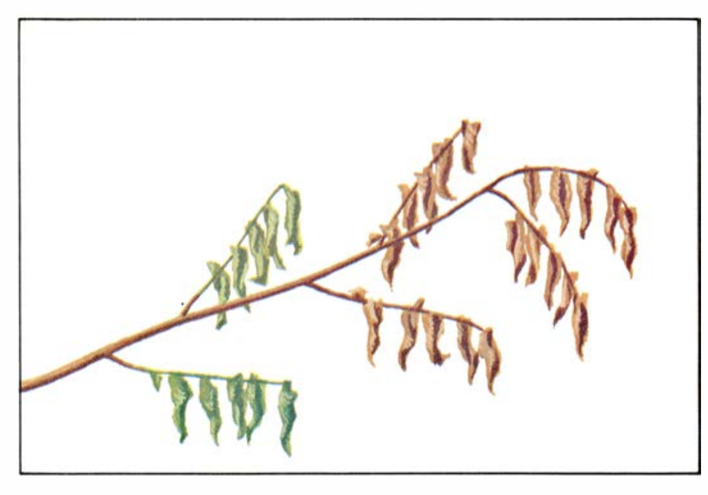
Los síntomas de la enfermedad holandesa del olmo comienzan con el marchitamiento de las hojas, normalmente en las puntas de las ramas. Las hojas, a la larga, pueden volverse amarillas, abarquillarse y, finalmente, adquirir una coloración marrón y morir. En otros casos pueden secarse rápidamente y tomar una coloración verde mate antes de morir. Cuando se pela la corteza de una rama infectada se aprecian claramente vetas marrones en la superficie de la madera. Aunque otros patógenos del olmo provocan algunos de estos síntomas, raramente son letales. (Una excepción es la necrosis del floema del olmo, o enfermedad amarilla del olmo, mucho menos frecuente que el mal holandés, aunque, al parecer, de consecuencias siempre fatales; se caracteriza por el amarilleamiento uniforme y muerte de toda la copa del árbol, a diferencia de la marchitez desigual típica de la enfermedad holandesa del olmo.)

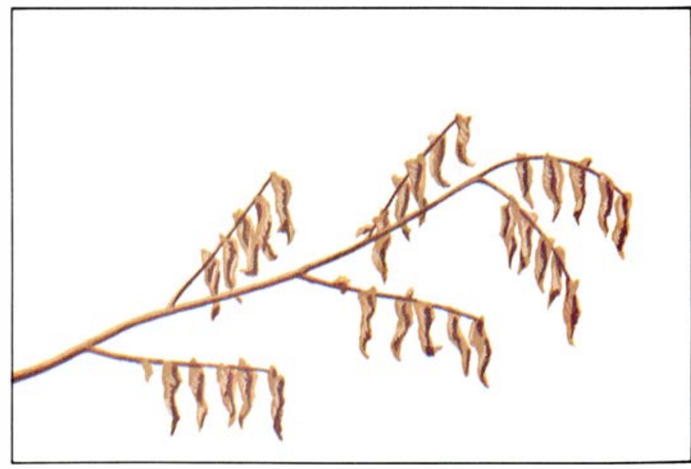
La expresión de los síntomas de la enfermedad holandesa del olmo varía con las distintas especies, época del año, el clima, el mecanismo de infección y la condición del árbol. El huésped principal en América del Norte es el olmo americano (*Ulmus americana*), pero la enfermedad es letal para especies menos comunes, entre ellas, el llorón (*U. rubra*), el de roca (*U. thomasii*), el olmo de septiembre (*U. serotina*) y el alado (*U. alata*). La mayoría de las especies europeas, incluidas el olmo blanco europeo (*U. laevis*), el olmo inglés (*U. procera*) y diversas variedades de *U. carpinifolia*, también son susceptibles. Por otra parte, muchas especies

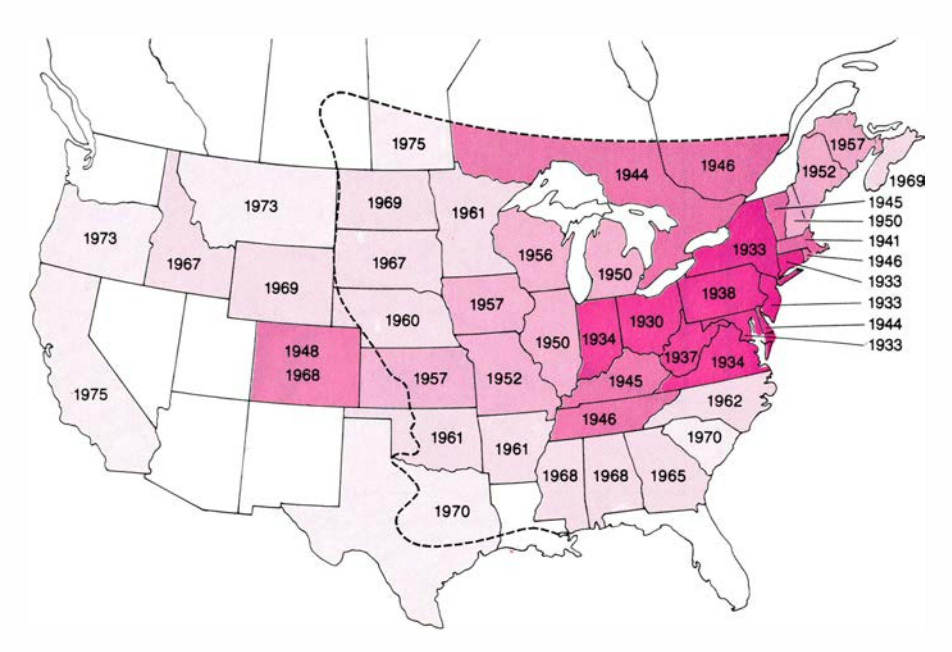
SINTOMAS CAUSADOS POR LA ENFERMEDAD HOLANDESA DEL OLMO, reflejados en las hojas de las ramas que se muestran en las dos series verticales de dibujos de la página opuesta. En ambos casos, las hojas se abarquillan o marchitan en las puntas de las ramas infectadas. En la serie de la izquierda, las hojas amarillean gradualmente y se abarquillan hacia dentro, antes de volverse marrones y morir. En la serie de la derecha, las hojas se secan rápidamente y toman un color verde mate antes de morir. En el primer caso, la mayoría de las hojas caen de la rama en cuanto mueren; en el segundo, pueden permanecer unidas a la rama varias semanas. Además de los síntomas foliares que en la ilustración se muestran, se pueden encontrar vetas marrones en la superficie de la madera cuando se pela la corteza.









DISPERSION de la enfermedad holandesa del olmo en Norteamérica. Las fechas señalan la aparición de la primera infección en cada estado y provincia. (Para Colorado se dan dos fechas porque, según parece, la enfermedad estuvo ausente en un intervalo de 20 años.) La gradación de los colores muestra el avance de la enfermedad a través del continente. La línea a trazos limita el área natural de *Ulmus americana*.

asiáticas, tales como el olmo chino (*U. parvifolia*) y el de Siberia (*U. pumila*), son, en comparación, resistentes a la infección.

Los síntomas que aparecen durante las primeras cuatro semanas tras la salida de las hojas en la primavera indican una infección residual o una infección contraída durante la época de letargo a través de las raíces injertadas en las raíces de un árbol enfermo. Estos síntomas tempranos afectan normalmente a más de un 10 por ciento de la copa del árbol. Una infección nueva típica, a partir de esporas transmitidas por el escarabajo de la corteza, surge en una sola rama a los 40 días o más después de salir las hojas. Un clima seco y cálido a principios del verano aumenta la marchitez, pero si la sequía dura mucho parece retrasarse la sucesión de los síntomas. Una infección avanza más aprisa y se manifiesta con mayor nitidez en árboles jóvenes y en los maduros vigorosos. En un olmo viejo, la infección nueva progresa lentamente; por esta razón, y también porque una infección transmitida por escarabajos de la corteza normalmente aparece primero en lo alto de la copa, la enfermedad holandesa del olmo no suele diagnosticarse en árboles viejos, hasta el segundo o tercer año de iniciado el brote infeccioso.

Por regla general, sólo los árboles infectados a través de las raíces mueren en el primer año de ataque. Un árbol infectado en primavera por escarabajos que se alimentan de sus ramas tiernas pierde parte de la cúpula en el otoño, y los síntomas progresan durante el año siguiente. Pero el infectado en mitad o a finales del verano puede escaparse al diagnóstico al desarrollarse la coloración normal del otoño. Ahora bien, en la primavera siguiente, los órganos afectados morirán y la infección se extenderá por todo el árbol.

Las barreras oceánicas contra la difusión de la enfermedad holandesa del olmo se superaron a principios de siglo, a raíz del comercio de troncos de olmo entre Europa central y Norteamérica. Los Estados Unidos importaban la madera para fabricar contrachapado. La enfermedad se había detectado en Inglaterra en 1927 y apareció en Cleveland, Ohio, en 1930. Por desgracia nadie hizo caso de las advertencias hechas por biólogos americanos en el sentido de prohibir la importación de troncos de olmo.

La infección de Ohio logró, aparentemente, erradicarse, pero persistieron centros independientes de la enfermedad en el área metropolitana de la ciudad de Nueva York, Nueva Jersey y el sureste de Connecticut, que se extendieron a pesar de los grandes y costosos esfuerzos de erradicación de la Oficina de Entomología y Cuarentena Vegetal. En 1944, la enfermedad se descubrió en el sur de Quebec y, a mediados de la década de 1970, había alcanzado la costa oeste del continente. Historias similares pueden relatarse de su avance a través de Europa y Asia occidental.

La enfermedad holandesa del olmo se ha descrito como una enfermedad de tipo epidémico, porque se extiende y crece rápidamente en una población de huéspedes susceptibles. La transmisión de la enfermedad se relaciona directamente con la distribución de los escarabajos vectores. Si bien los escarabajos pueden, por sí mismos, dañar los árboles, ello raramente acarrea la muerte del olmo. Aunque el escarabajo europeo de la corteza del olmo precedió al hongo en Norteamérica en dos décadas, al menos, las especies norteamericanas no estaban amenazadas de gravedad hasta que llegó el hongo. Una vez coincidieron escarabajo, hongo y una gran población de huéspedes susceptibles, brotó la epidemia.

El principal factor que determina la tasa de infección de la enfermedad holandesa del olmo es el tamaño de la población de escarabajos, que, a su vez, depende de la disponibilidad de tejido vegetal adecuado para la reproducción. Un olmo sano puede resistir con éxito los intentos de crianza de los escarabajos de la corteza de olmo; los escarabajos tienden a reproducirse en árboles afectados por la sequía, heridos o enfermos. En consecuencia, el hongo aumenta la población de escarabajos al ampliar el hábitat de cría, y la creciente población de artrópodos provoca un aumento masivo de nuevas infecciones. La etapa crucial de la epidemia es la primera parte. A menos que la infección se controle al principio, la población de escarabajos y de hongos se multiplicará rápidamente, lo que resultará en una difusión exponencial de la epidemia.

Ceratocystis ulmi, el hongo que causa la enfermedad holandesa del olmo, tiene dos etapas reproductivas en su ciclo biológico: una etapa es asexual y la otra sexual. Las esporas infecciosas se producen en las dos etapas. Las estructuras asexuales esporuladas, llamadas sinemas, se caracterizan por un pedúnculo marrón azulado de aproximadamente un milímetro de largo. El pie es un haz de filamentos, llamados hifas, que se ensanchan al final para formar una cabeza incolora, donde se encuentran un gran número de esporas, suspendidas en una gotita mucilaginosa. Los sinemas se forman rápidamente en los árboles infectados, especialmente en las galerías o túneles horadados por el escarabajo de la corteza, facilitando así el contacto entre el vector y el patógeno.

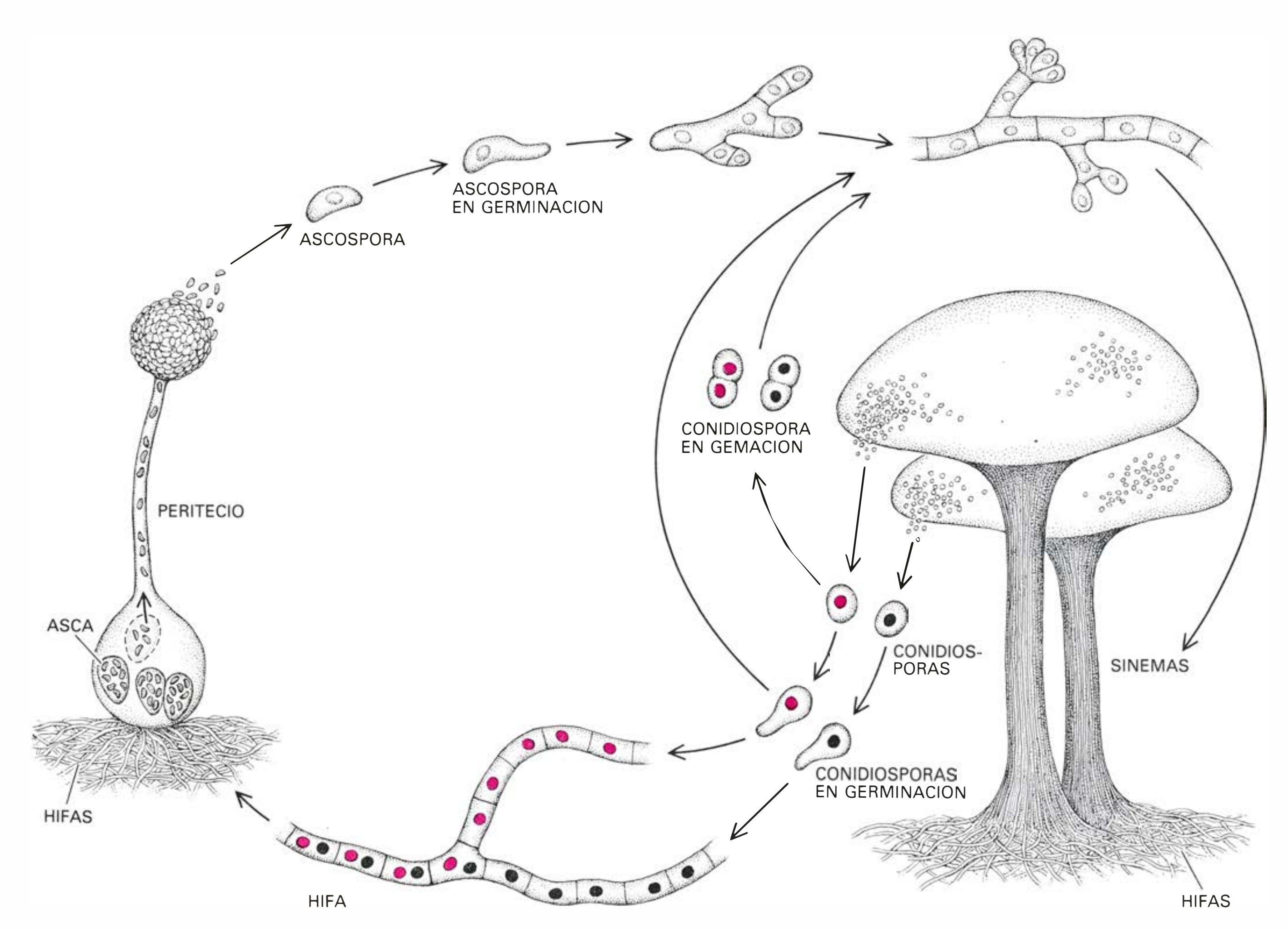
La forma sexual del hongo, llamado peritecio, se halla frecuentemente en las mismas galerías. Su configuración semeja la de una botella con cuello largo, y aparece después de la unión de dos formas sexuadas del hongo. Las esporas crecen en el peritecio en grupos

de ocho, encerrados en una envuelta denominada asca, que se desarrolla después de completarse el proceso sexual de fusión nuclear y división celular. Hay muchas ascas en cada peritecio. Cuando se liberan las esporas de las ascas, flotan libremente en la parte esférica del peritecio antes de ser expulsadas a través de un orificio.

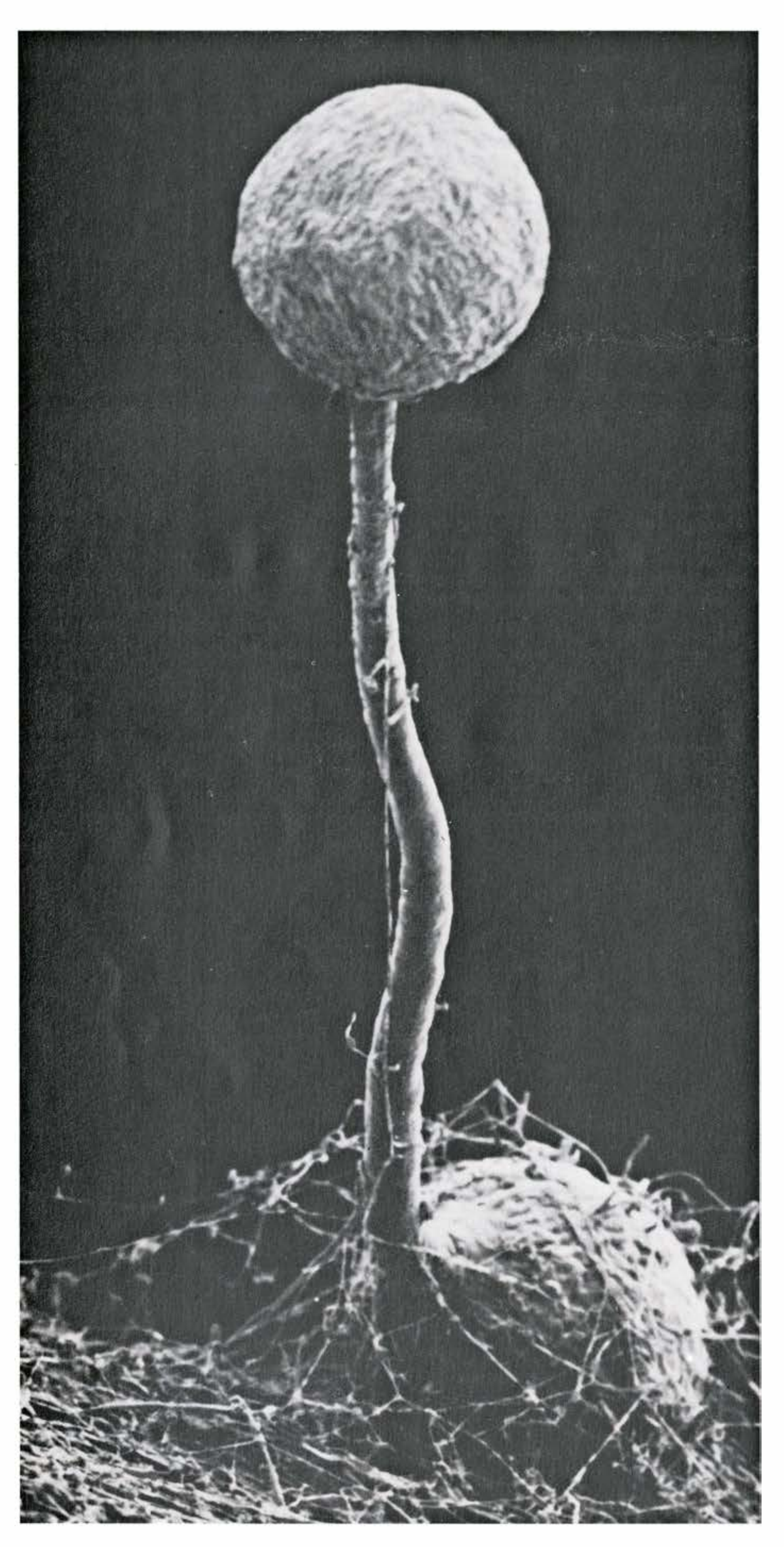
Desde el punto de vista bioquímico, hay más interrogantes que respuestas en el morbo holandés. ¿Por qué *Ceratocystis ulmi* provoca los síntomas de la enfermedad sólo en unas pocas especies de árboles (los olmos susceptibles de América y Europa) y no en otras plantas? ¿Acaso depende el crecimiento del hongo de alguna sustancia nutritiva que se encuentra exclusivamente en olmos vivos? No parece que sea así. El hongo puede crecer sobre madera de olmo muerto, sobre medio de cultivo preparado a partir de extractos de vegetales e incluso sobre un medio sintético. Pero cuando lo inyectamos en otro vegetal, un arce o un fresno, apenas si crece, no causa ningún daño apreciable y pronto se muere.

Si el crecimiento del hongo en los olmos susceptibles no depende de una sustancia nutritiva, ¿puede relacionarse con la síntesis de una sustancia promotora del crecimiento, particularmente en mayo y junio, cuando los árboles se muestran más susceptibles a la infección? Quizá haya moléculas de "reconocimiento" en el hongo y en el huésped. Pero, ¿qué serán? Tales preguntas carecen todavía de respuesta.

Acabamos de asistir a cierto avance en nuestro conocimiento del mecanismo mediante el cual el hongo daña al árbol: parece que el hongo secreta una toxina. Se ha demostrado experimentalmente que hay una sustancia que causa la marchitez de los olmos; esta sustancia ha sido aislada por un equipo de investigadores, entre los que se encuentran Kenneth D. Hapner y uno de los autores (Strobel), de la Universidad estatal de Montana, Peter Albersheim, de la Universidad de Colorado en Boulder, Neal K. Van Alfen, de la estatal de Utah, y John H. Nordin, de la de Massachusetts en Amherst. La existencia de una toxina había sido postulada previamente sobre bases teóricas, pero no podía caracterizarse hasta aislarla de entre las muchas sustancias que hay en los cultivos de hongos. Hemos demostrado que la sustancia tóxica es, en realidad, una mezcla de moléculas, que reciben el nombre colectivo de peptidorhamnomanana. Cada molécula consta de un esqueleto proteico con, al menos, dos o tres cadenas laterales polisacáridas, esto es, cadenas con muchas unidades de azúcar. El esqueleto proteico está enriquecido en los aminoácidos treonina y serina, a los que están unidas las cadenas polisacáridas [*véase la ilustración de la página 40*]. Además, unidas irregularmente sobre la proteína hay cadenas más cortas de unas pocas unidades de azúcares. Las mediciones del peso molecular de la toxina dan resultados variables debido a



CICLO BIOLOGICO de *Ceratocystis ulmi*, el hongo causante de la enfermedad holandesa del olmo. Todas las formas por las que pasa el organismo a lo largo de su ciclo biológico se pueden encontrar en las galerías perforadas por los escarabajos de la corteza del olmo. El hongo tiene dos cuerpos fructíferos, uno sexual, llamado peritecio y otro asexual, o sinema. El peritecio consta de numerosas estructuras, ascas, cada una de las cuales encierra ocho ascosporas. Cuando las hifas, o filamentos, derivados de las conidiosporas de dos razas de *C. ulmi* sexualmente compatibles se funden, se unen los dos tipos de núcleos celulares (*puntos negros y de color*) y se forma un peritecio, donde tiene lugar la meiosis; su resultado es la producción de ascosporas, en las que se ha dado una redistribución de los cromosomas. Conidiosporas, ascosporas y las esporas producidas por gemación (*arriba, derecha*) son infecciosas.



CUERPO FRUCTIFERO SEXUAL, o peritecio, del hongo productor de la enfermedad holandesa del olmo, según una micrografía electrónica de barrido tomada recientemente por Ruud J. Scheffer, del Willie Comelin Scholten Laboratory de Baarn, Holanda. La estructura esférica situada en lo alto del pie de hifas está repleta de ascosporas infecciosas. El aumento es, aproximadamente, de 3000 diámetros.

las variaciones en la longitud de las cadenas laterales polisacáridas. El peso molecular medio es muy alto, alrededor de 130.000.

Van Alfen, en colaboración con Neil Turner, de la Estación Experimental Agrícola de Connecticut, demostraron que la toxina peptidorhamnomanana no dañaba directamente las células de un olmo cual si actuara de veneno. Por el contrario, parecía interferir el movimiento del agua a través del xilema, o núcleo leñoso del árbol. La interferencia parecía tener lugar en las estructuras llamadas punteaduras, de la pared de los vasos del xilema. El aporte de agua a las hojas se mantiene gracias al flujo que existe desde los vasos de la rama hasta los vasos de la hoja, a través de las punteaduras. Si el agua se transpira en la superficie de las hojas más de prisa de lo que la suministra el árbol, las hojas y tallos se marchitan. Si persiste el déficit hídrico, los orgánulos celulares, como los cloroplastos verdes, resultan dañados y las hojas se vuelven amarillas. Estos efectos se observan cuando se administra la toxina peptidorhamnomanana a esquejes de olmo. Hay otras macromoléculas con el mismo efecto, si bien no se localizó ninguna en cultivos de *Ceratocystis ulmi*.

Ruud J. Scheffer y Doekle M. Elgersma, del laboratorio Willie Comelin Scholten, de Holanda, trabajando con técnicas inmunológicas muy sensibles, comprobaron la presencia de pequeñas cantidades de peptidorhamnomanana en olmos enfermos, pero no en los sanos. Puesto que cantidades mínimas de la toxina pueden reducir de forma apreciable el flujo de agua en esquejes de olmo, parece que la toxina está presente en los tejidos enfermos en proporción suficiente para causar daño. Sin duda, la sustancia desempeña algún papel en el avance natural de la enfermedad, aunque no podamos afirmar que la peptidorhamnomanana sea lo que determina la especificidad del hongo *C. ulmi* hacia el olmo. Tal vez se hallen implicados otros factores.

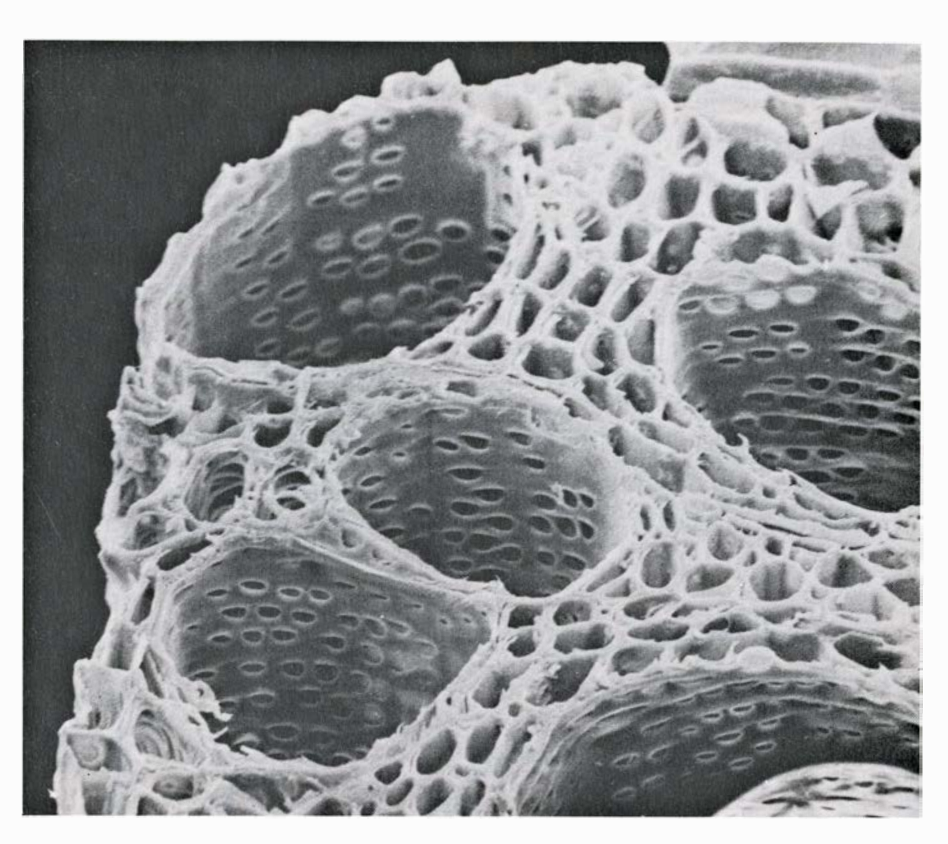
Sin un insecto vector, la enfermedad holandesa del olmo no alcanzaría la magnitud que hoy tiene. La razón estriba en que el diseño de las esporas de *C. ulmi* no se adapta bien a la dispersión por el viento, pero sí a la transmisión por insectos. Muchas especies de insectos pueden contaminarse con esporas fúngicas, pero sólo los escarabajos de la corteza son vectores normales de la enfermedad, pues son los únicos que se mueven de árboles enfermos a árboles sanos.

Los escarabajos de la corteza del olmo crían en el floema de árboles debilitados, enfermos, moribundos o recién cortados. Las hembras dejan los huevos a lo largo de las paredes de las galerías de puesta. Hacia los ocho días, la larva sale del huevo y perfora túneles de alimentación que terminan en la cámara pupal. Tras la maduración de las pupas, los adultos perforan una salida en la corteza y se dispersan, normalmente volando a otro olmo debilitado, para criar, o a un olmo sano, para alimentarse. Los escarabajos suelen alimentarse en la horcadura de las ramitas o en grietas existentes en la corteza de ramas jóvenes. Mientras comen, los escarabajos practican surcos en la madera e infectan los olmos sanos.

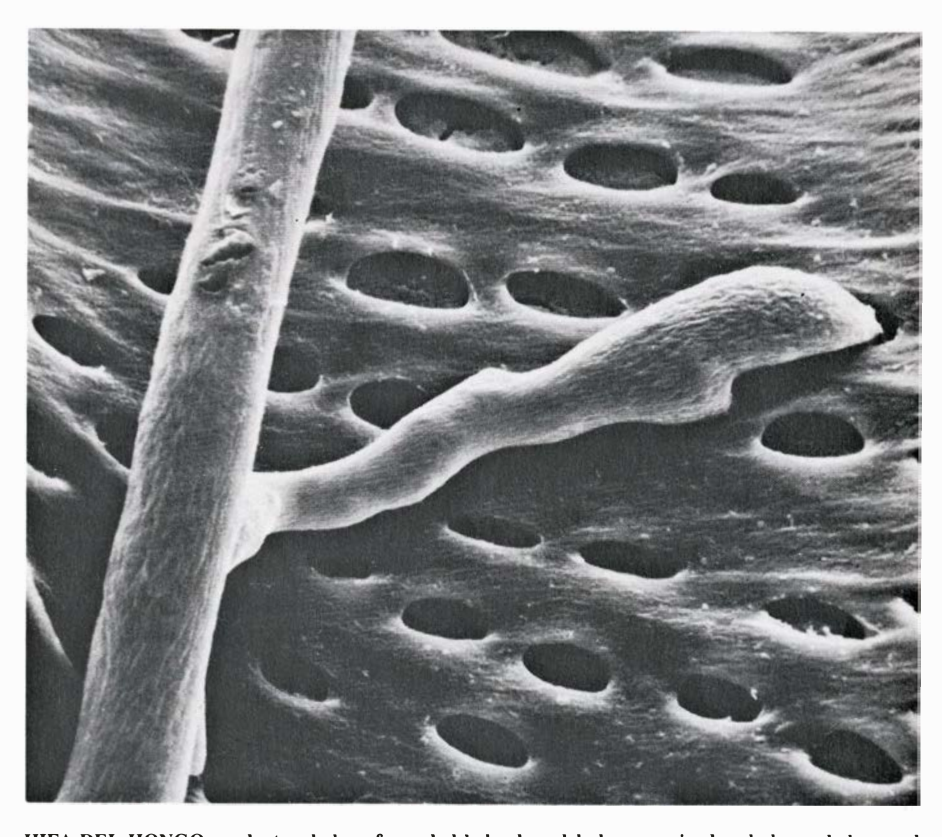
En Europa, los dos vectores principales de la enfermedad holandesa del olmo son el pequeño escarabajo de la corteza del olmo (*Scolytus multistriatus*) y la especie más grande *S. scolytus*. Otras especies de *Scolytus* que crían sobre los olmos más dos del género *Pteleobius* se citan también como transmisores de la enfermedad. En Norteamérica, los principales vectores son un escarabajo de la corteza indígena, *Hylurgopinus rufines*, y *S. multistriatus*, que suele denominarse escarabajo europeo de la corteza, por haberse introducido en el este de los Estados Unidos procedente de Europa poco antes de 1909.

Los escarabajos de la corteza europeos e indígenas son similares en tamaño (entre dos y 3,5 milímetros de longitud) aunque de morfología distinta. La especie europea, brillante, presenta dos tonalidades: de color pardorrojizo oscuro los élitros y, negro, el tórax. De un gris parduzco mate, el escarabajo nativo posee una textura áspera. Las etapas inmaduras de los dos tipos pueden distinguirse fácilmente por el sistema de galerías en que se encuentran. La hembra europea taladra la galería de puesta a lo largo de la fibra de la madera, y los túneles que excava la larva irradian de la galería de puesta, resultando una forma elíptica. La galería de puesta del escarabajo indígena atraviesa la fibra, y la larva tiende a perforar a lo largo de la fibra, creando un dibujo en forma de mariposa.

El escarabajo indígena se halla en todo el área natural de distribución del olmo en Norteamérica; no se ha localizado, empero, en los estados de las Montañas Rocosas y en el oeste, donde la introducción de los olmos ha sido más reciente. El escarabajo europeo se encuentra allí donde hay olmos en Norteamérica, salvo en puntos del norte



SECCION TRANSVERSA de madera de olmo aumentada 700 diámetros en una micrografía electrónica de barrido obtenida por Doekle M. Elgersma, colega de Scheffer. Los grandes vasos que corren longitudinalmente a lo largo del xilema son los conductos principales por los que circula el agua en el árbol. Cada punteadura, u orificio, de las paredes de los vasos del xilema tiene una tenue membrana que permite el paso del agua a su través en perpendicular al vaso. Las membranas no alcanzan a verse aquí.



HIFA DEL HONGO productor de la enfermedad holandesa del olmo, creciendo a lo largo de la pared interior de un vaso del xilema de un trozo de madera de olmo. Una rama lateral se destaca de la hifa principal y penetra en dos de las punteaduras. Las membranas de las punteaduras, algunas de las cuales llegan a verse a este aumento, aparecen como platos cascados. Se cree que una sustancia tóxica segregada por *Ceratocystis ulmi* interfiere el movimiento del agua a través de la membrana de las punteaduras. La micrografía electrónica de barrido es de Elgersma. El aumento, en este caso, es de unos 2500 diámetros.

donde las temperaturas inferiores a -4 grados Celsius matan las larvas invernales.

Ambos escarabajos son comunes en el este de los Estados Unidos. Podemos distinguirlos allí por sus hábitos de cría. Las especies europeas, más agresivas, atacan los árboles en las primeras etapas de su decadencia, mientras que los escarabajos nativos colonizan tejido moribundo. En general, el escarabajo europeo monopoliza árboles que crecen aislados y la parte superior soleada de las copas de los árboles inmediatos. La evitación, por parte del escarabajo europeo, de áreas sombreadas permite al escarabajo indígena prosperar en parques, bosquecillos y bosques. El europeo suele iniciar su ataque en lo alto de la copa de un árbol enfermo, mientras que el escarabajo nativo se concentra en las ramas inferiores y el tronco. Si las dos especies atacan al mismo árbol, los escarabajos nativos son superados pronto por la descendencia de los europeos, que se desarrolla más de prisa.

La rápida maduración del escarabajo europeo le permite completar dos generaciones por año en la parte norte de su área de distribución. En las zonas del sur comienza una tercera generación, donde a veces concluye también. Los escarabajos invernan en el interior de las galerías en forma de larvas completamente desarrolladas, para abandonarlas, adultos ya, entre primeros de abril y mediados de junio.

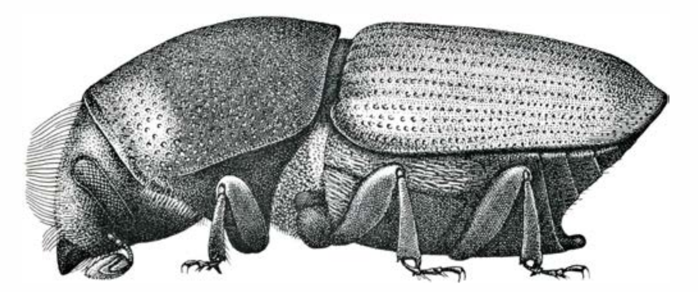
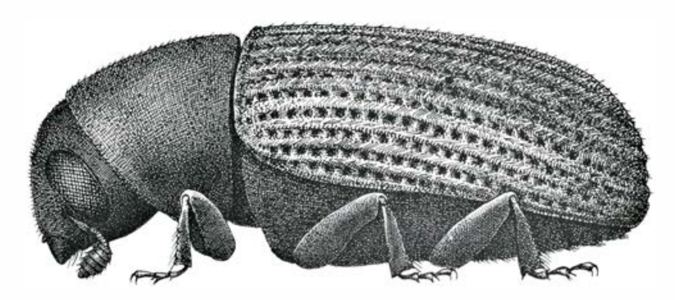
El escarabajo de la corteza indígena completa sólo una generación por año (o a lo sumo una y media) e inverna en forma de neoadulto o de larva. Los adultos que van a invernar salen del lugar de cría al final del verano y vuelan a un olmo vivo, donde se alimentan de la corteza de las ramas. Al aproximarse el otoño, descienden para cortar cámaras de hibernación en el corcho, o corteza exterior, de la parte inferior del tronco. Al inicio de la primavera, los túneles se extienden en nichos de alimentación a lo largo del floema. Durante abril y mayo, los adultos invernantes vuelan a árboles moribundos, donde procrean. Algunos de los adultos abandonan la progenie en mayo y junio para producir una nueva generación en otro árbol. Los adultos de la nueva nidada emergen del árbol en verano y los primeros producen crías que pasarán el invierno siguiente como larvas.

Un olmo sano es más susceptible a la infección cuando está formando madera nueva, entre finales de mayo y primeros de junio. Por tanto, los escarabajos europeos que se hacen adultos en primavera y los escarabajos nativos que invernan adultos muestran una notable eficacia en la transmisión del hongo. El escarabajo de la corteza europeo extiende la enfermedad cuando los adultos que emergen vuelan a un olmo sano y perforan las horcaduras de las ramas, donde depositan el hongo. El escarabajo indígena probablemente infecte olmos sanos en la primavera, cuando prolonga sus túneles de invernación hacia la madera de su huésped o asciende por el árbol y se alimenta de las ramas.

Aunque los escarabajo pueden colonizar diferentes partes de un árbol durante dos o más generaciones, la corteza de una parte del mismo sólo se utiliza una vez para la reproducción; de ahí que cada generación deba encontrar un nuevo lugar de puesta. El proceso de localizar corteza moribunda para reproducirse y madera sana para alimentarse e invernar está regulado por aromas químicos liberados por el árbol huésped y por los escarabajos.

ESTRUCTURA MOLECULAR de la peptidorhamnomanana, una sustancia tóxica aislada de cultivos del hongo productor de la enfermedad holandesa del olmo. La representación esquemática se basa en un modelo construido por John H. Nordin, de la Universidad de Massachusetts en Amherst y uno de los autores (Strobel). La molécula tiene un esqueleto de proteína formado por una larga cadena de aminoácidos (*puntos negros*). Unidas a la cadena proteica central hay varias cadenas largas de polisacáridos. Cada cadena de polisacáridos consta de unidades de azúcar manosa unidas entre sí (*puntos coloreados*) a las que se juntan numerosas unidades terminales del azúcar ramnosa (*círculos vacíos*). Las cadenas laterales más cortas están compuestas de unas pocas unidades de azúcar y se unen al azar sobre la proteína. El peso molecular de la toxina varía en función de la longitud de las distintas cadenas polisacarídicas laterales.

Durante la última década, uno de nosotros (Lanier) junto con otros investigadores de la Facultad de Ciencias del Medio Ambiente y Silvicultura de la Universidad del estado de Nueva York en Syracuse, han venido tra-

DOS ESPECIES DEL ESCARABAJO de la corteza del olmo son los principales vectores, o portadores, de la enfermedad holandesa del olmo en Norteamérica. La especie europea (*Scolytus multistriatus*) se representa a la izquierda por un individuo macho. Es brillante y presenta dos tonalidades: pardo rojizo oscuro en los élitros y negro en el tórax. La parte frontal de la cabeza está cubierta por un cepillo de pelos dorados llamados sedas. La especie indígena (*Hylurgopinus rufipes*) se representa a la derecha. Tiene un color pardo grisáceo mate y una textura más áspera. Las dos especies muestran siluetas claramente diferenciadas. El escarabajo europeo es cóncavo en la parte inferior posterior del abdomen, mientras que la especie nativa tiene un perfil más redondeado. Los escarabajos adultos de ambas especies son similares en su tamaño: entre dos y 3,5 milímetros. La ilustración es de Tom Prentiss.

bajando con un grupo del laboratorio de Arboles de Sombra, del Servicio Forestal de los Estados Unidos en Delaware, Ohio, en el estudio de la ecología química del escarabajo de la corteza europeo. Nos proponíamos reducir la tasa de la enfermedad holandesa del olmo a través de la modificación del comportamiento del artrópodo con aromas químicos; el objetivo se ha cubierto en parte.

En virtud de la radiación ultravioleta solar, los escarabajos de la corteza europeos que emergen se ven impulsados a dispersarse volando. Alcanzan vuelos de varios kilómetros. Remontado el vuelo, el escarabajo responde a los aromas de los olmos, particularmente a los de los árboles débiles o enfermos. Además, cuando una hembra encuentra un posible lugar de cría, libera un aroma, llamado feromona de agregación, que anuncia la presencia de madera adecuada para la colonización. En consecuencia, un árbol en el que haya aterrizado una sola hembra puede convertirse en el objetivo de una invasión masiva.



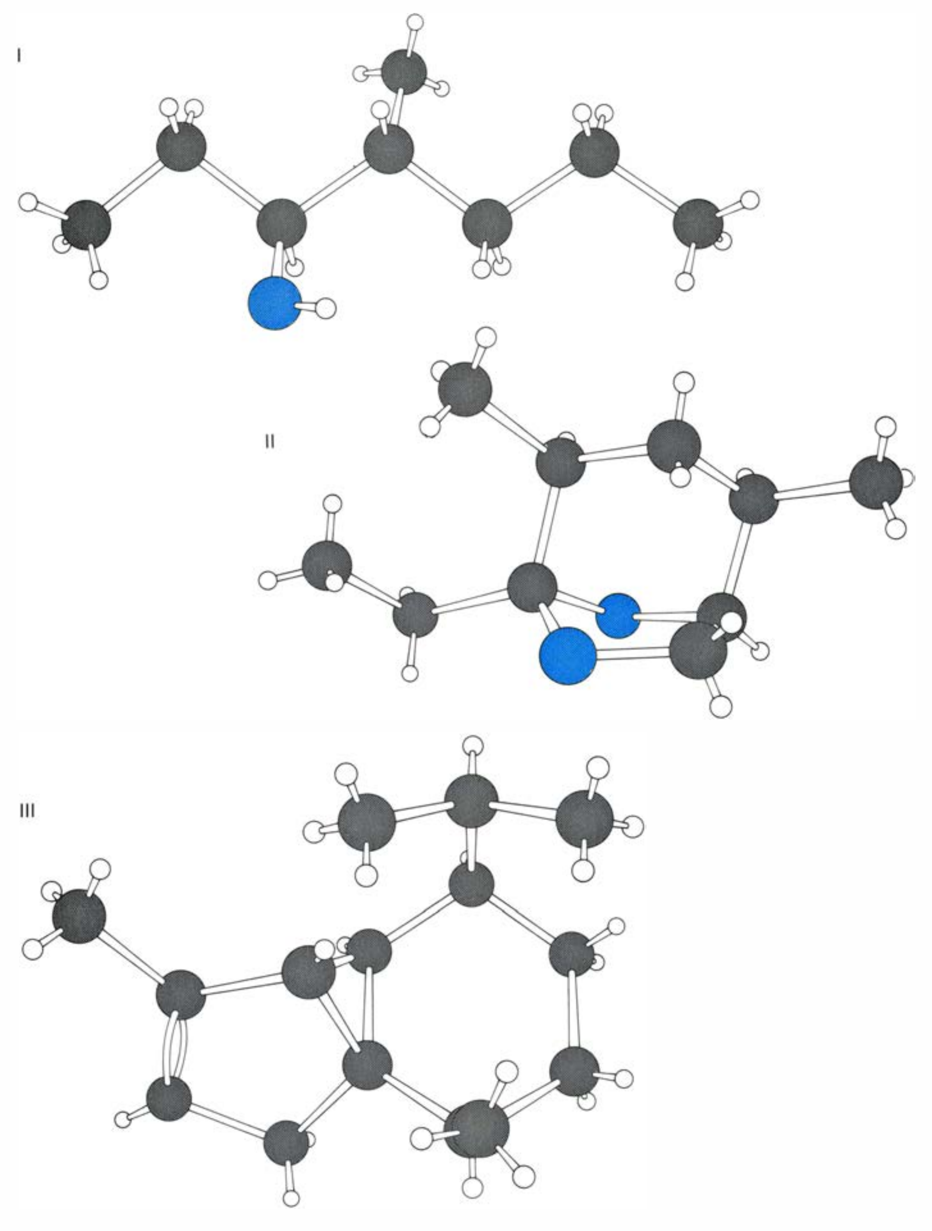

DIBUJO CARACTERISTICO, grabado en la madera de un árbol enfermo, por el escarabajo europeo de la corteza (*fotografía izquierda*). Los canales profundos que corren paralelos a la veta de la madera son las galerías de puesta, excavadas por hembras adultas. Las líneas más superficiales que irradian de cada una de las galerías de puesta son galerías de alimentación practicadas por las larvas. En la fotografía de la derecha se observa con gran aumento una hilera de larvas alimentándose en el floema, o corteza interior, de un olmo. A diferencia de los escarabajos de la especie europea, los escarabajos indígenas de la corteza del olmo suelen excavar sus galerías a través de la veta de la madera y las galerías de alimentación de las larvas paralelas a la veta.

Para aislar e identificar la feromona, se extrajeron sustancias volátiles de una corriente de aire que pasaba sobre troncos de olmos infectados con miles de hembras vírgenes del escarabajo europeo de la corteza del olmo. (Se excluyeron los machos porque nuestros estudios habían demostrado que el efecto atractivo de las feromonas cesaba tras el apareamiento.) El extracto se dividió en fracciones, recogiendo partes del mismo a intervalos de tiempo sucesivos a medida que pasaban por un cromatógrafo de gases. Los bioensayos de laboratorio, en que se ponía a los escarabajos frente a varias muestras hacia las que podían desplazarse, demostraron la atracción que ejercía el extracto total y posteriormente permitieron identificar fracciones que no podían eliminarse sin atenuar la respuesta. Las pruebas de campo confirmaron que la atracción de una fracción, por sí sola, era mínima, en tanto que una combinación de tres fracciones determinadas producía una fuerte respuesta.

Fraccionamientos repetidos y bioensayos de las tres fracciones permitieron obtener tres compuestos puros en cantidades de microgramos. Estos compuestos se identificaron por técnicas espectrofotométricas y fueron sintetizados por Robert M. Silverstein y sus colaboradores en Syracuse. Los nombres químicos son 4-metil-3-heptanol, (−)α-multistratina y (−)α-cubeno; para simplificar nos referiremos a ellos como los compuestos I, II y III, respectivamente [*véase la ilustración inferior*].

Los tres compuestos son necesarios para obtener un pico máximo de atracción del escarabajo. El compuesto III es un producto de la corteza de los olmos moribundos o infectados; el I y el II los liberan las hembras vírgenes del escarabajo. Después del apareamiento, se detiene la liberación del compuesto I y cesa la atracción. Si otros escarabajos se dirigen hacia el centro de agregación, son desviados por el exceso del compuesto II sobre el I hacia olmos moribundos próximos, donde crían, o hacia olmos sanos, donde se alimentan. Este sistema de comunicación química justifica la mayor parte del comportamiento colonizador de los escarabajos y la naturaleza contagiosa de la enfermedad holandesa del olmo.

La ecología química del escarabajo de la corteza del olmo indígena difiere de la de las especies europeas. El factor principal y probablemente único, en la atracción a larga distancia, son las sustancias volátiles de la corteza del olmo moribundo. Puesto que el escarabajo nativo normalmente ataca la madera de olmo en un estado de descomposición mucho más avanzado que la madera invadida por el escarabajo europeo, un posible huésped moribundo probablemente difiera de forma significativa de un árbol sano y de aquellos que, estando menos dañados, los prefieren los escarabajos europeos.



FEROMONAS DE AGREGACION, que logran atraer al escarabajo europeo de la corteza del olmo *Scolytus multistriatus*, y que se han identificado, mediante fraccionamientos repetidos y bioensayos, y sintetizado. En estos diagramas estructurales los átomos de carbono son negros, los átomos de hidrógeno blancos y, los de oxígeno, de color. De arriba abajo aparecen el 4-metil-3-heptanol, la (−)α-multistratina y el (−)α-cubebeno; por mor de simplificación, se les denomina compuestos I, II y III, respectivamente. Los compuestos I y II los producen los escarabajos, mientras que el III emana de los olmos heridos.

Se han diseñado cinco estrategias principales para controlar la enfermedad holandesa del olmo. En primer lugar, la cuarentena de ciertas áreas permitirá su salvación. Segundo, la enfermedad puede erradicarse mediante la tala de árboles enfermos irrecuperables y la eliminación de madera de olmos que probablemente vaya a ser colonizada por los escarabajos (combinación de medidas que toman el nombre global de saneamiento). Tercero, puede reducirse la tasa de infección de una región controlando la tasa de reproducción, ya sea del hongo o del escarabajo. Cuarto, los árboles de particular valor podrán preservarse mediante varios métodos, incluida la aplicación de insecticidas y fungicidas. Quinto, las especies de olmo sensibles al morbo pueden reemplazarse por otras resistentes. Cabe también la combinación de dos o más estrategias de éstas.

La cuarentena y erradicación sólo

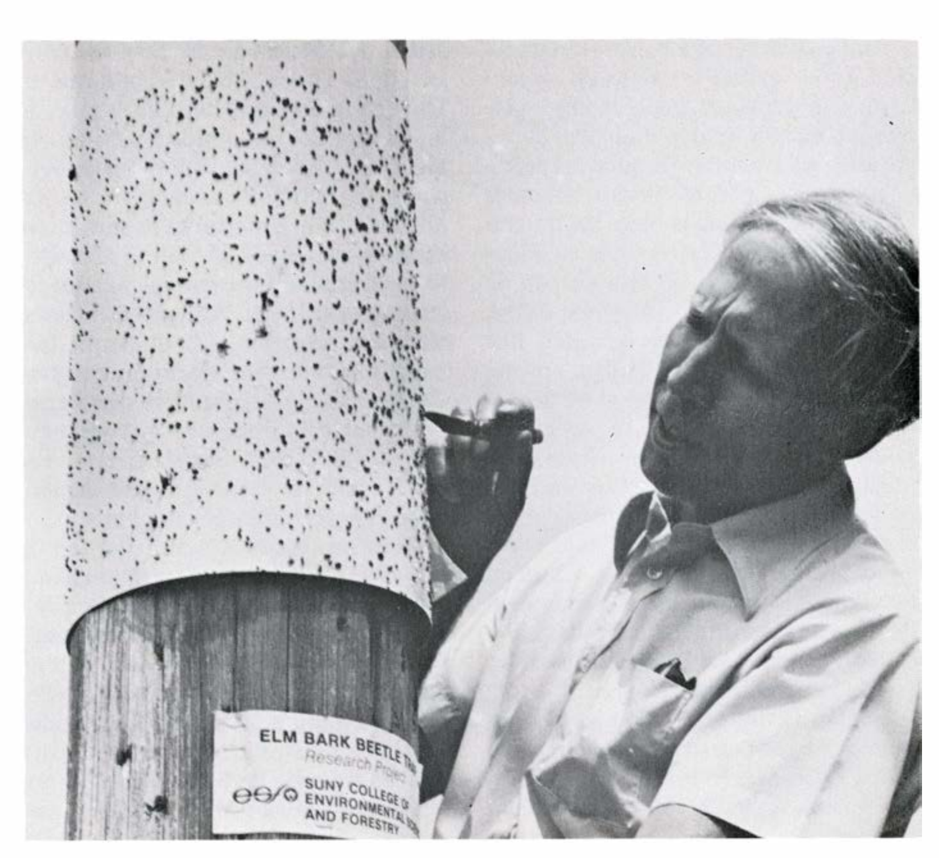
son viables en las zonas donde no haya llegado la enfermedad o sean contados los casos identificados. Se intentó la erradicación en el este de la Unión antes de la segunda guerra mundial, pero resultó un fracaso, porque se había minusvalorado la capacidad de vuelo del escarabajo de la corteza. Se pensó, también, que la madera cortada no intervenía en el proceso de infección; además, abundaban los olmos silvestres, en los que la enfermedad podía intensificarse sin que nadie lo controlara o advirtiera. El control práctico de la enfermedad, cuando se ha conseguido, siguió al dominio del hongo, o del escarabajo, y, en menor grado, al tratamiento preventivo y terapéutico de los árboles, uno a uno.

Las estrategias más rentables en el control de una plaga dependen de varios factores, que van desde el tamaño de la población de escarabajos hasta el número de olmos, pasando por su densidad y el valor asignado a los árboles. Pero hay una medida imprescindible en cualquier programa de control: los árboles enfermos deben tratarse o talarse antes de que se constituyan en fuente de nuevas crías de escarabajos.

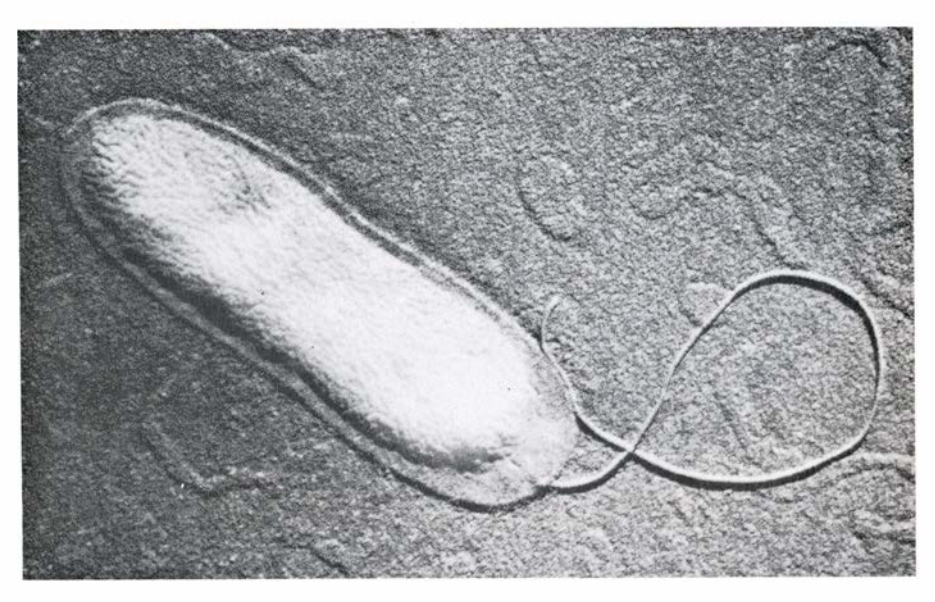
La protección de un determinado olmo, en ausencia de medidas eficaces de saneamiento en la población de los alrededores, no es muy prometedora. El coste de un tratamiento preventivo anual, ya sea pulverizando insecticida o mediante inyecciones de fungicida, puede superar los 100 dólares para un solo árbol de gran tamaño. Mas si la enfermedad se halla asentada y abundan los escarabajos de la corteza, ni siquiera con los propios tratamientos anuales se logra más allá de una protección muy limitada. Ningún método en particular, ni la combinación de varios, puede resultar capaz de mantener bajos niveles de la enfermedad holandesa del olmo, si no se controla el suministro de material donde prolifere el escarabajo.

Entre los nuevos métodos para controlar la enfermedad holandesa del olmo están los que pretenden manipular el comportamiento de los escarabajos con feromonas y sustancias que atraen al huésped y atacar, biológicamente, al hongo con un organismo antagonista. Se están ensayando planes para una estrategia integral de control dirigida contra los escarabajos y el hongo.

Desde 1975, el grupo SUNY, dirigido por uno de nosotros (Lanier), ha desplegado trampas pegajosas, cebadas con una mezcla artificial de feromonas del escarabajo europeo de la corteza del olmo, que ciñen pequeños bosques de este árbol en 12 puntos de ocho estados. Tales olmos recibían ya cuidados, pero continuaban sufriendo nuevas infecciones del mal holandés a una tasa anual cifrada en el 5 y el 20 por ciento de la población sana. En general, el saneamiento de los bosques era excelente. En relación con esto, un grupo del Servicio Forestal con base en Ohio, y dirigido por John W. Peacock, había hecho un experimento a gran escala para atrapar escarabajos en Fort Collins, Colorado y Evanston, Illinois.



ESCARABAJOS ATRAPADOS sometidos a examen por uno de los autores (Lanier) durante un experimento realizado en la Facultad de Ciencias del Medio Ambiente y Silvicultura de la Universidad del estado de Nueva York en Syracuse. Fueron atraídos al papel engomado por una feromona de agregación sintética. Los insectos más pequeños son escarabajos europeos de la corteza del olmo de ambos sexos.



BACTERIA sometida al ensayo de su eficacia para el tratamiento de la enfermedad holandesa del olmo en árboles aislados. La bacteria, de la especie *Pseudomonas syringae*, tiene un apéndice en forma de látigo, llamado flagelo, que le permite moverse. Sintetiza productos antifúngicos. La micrografía es de Bruce Hemming, de la Universidad estatal de Montana. El aumento, aproximadamente, de 40.000 diámetros.

La incidencia sobre la tasa de infección de la captura en masa de escarabajos europeos, realizada en el programa que incluía las dos ciudades, ha resultado ambivalente. Aunque las pérdidas causadas por el morbo holandés eran, en general, más bajas en las ciudades ensayadas que en otras ciudades de la misma región, la tasa de tala de árboles no era inferior a lo que cabría esperar sin la caza de escarabajos. Los bosques tomados como testigo, protegidos con trampas cebadas con feromona, dieron resultados más prometedores. Con pocas excepciones, la tasa de infección disminuyó en todos los años de tratamiento. Cuando se interrumpió el programa de trampas en algunos bosquecillos, en 1978, la tasa de infección aumentó en todos los casos, excepto en uno, mientras continuaba remitiendo o se mantenía a cero en los bosquecillos donde persistían las trampas.

¿Acaso la destrucción de padres potenciales puede causar un gran impacto sobre el tamaño de la siguiente generación? Experimentos con trampas realizados por Martin C. Birch y sus colegas, de la Universidad de California en Davis, sugieren que no. Los aromas de la madera del olmo moribundo atraen algunos escarabajos incluso en presencia de la feromona artificial. Una vez que las hembras perforan la corteza, la madera, debilitada por la colonización, se hace competitiva con las fuentes de feromona artificial. Aunque se recluten menos escarabajos para el ataque que los que habría que esperar en ausencia de trampas, los ejemplares que llegan del árbol crean galerías más largas y con más huevos; la propia supervivencia de las larvas es mayor, dada la menguada competencia por el alimento.

Nuestros mejores resultados en la protección de olmos sanos mediante trampas pegajosas cebadas con feromonas se consiguieron en áreas comparativamente libres de madera de cría y de material recién infectado que compitiera con la feromona. Por desgracia, tales áreas son raras. Parece claro que, para que las trampas de feromona alcancen resultados eficaces contra *Scolytus multistriatus*, deben intensificarse los esfuerzos de saneamiento. Los obstáculos fiscales y políticos que se oponen a una mejora de las prácticas de saneamiento a gran escala por medio de métodos tradicionales son formidables. Se necesitan, por tanto, nuevas técnicas de saneamiento.

INHIBICION del hongo *Ceratocystis ulmi*, causante de la enfermedad holandesa del olmo, por la bacteria antagonista *Pseudomonas syringae*. Se aprecia claramente en este ensayo de laboratorio. Cuatro colonias de *P. syringae* se sembraron sobre una placa de Petri; en la fotografía corresponden a los cuatro puntos, esquinas de un cuadrado, con coloración crema. Sobre la placa se pulverizó una suspensión de esporas y fragmentos de hifas de *C. ulmi*. Tras varios días de incubación, el hongo podía ya distinguirse en forma de una lámina que recubría la superficie de la placa, excepto las zonas circulares alrededor de las colonias de bacterias. Un agente antimicótico sintetizado por la bacteria inhibe el crecimiento del hongo.

El hongo de la enfermedad holandesa del olmo no es necesario para generar madera de olmo que atraiga a los escarabajos de la corteza. Por ejemplo, olmos próximos a la muerte también atraen al escarabajo indígena tras exponerlos al herbicida ácido cacodílico.

Nos hemos aprovechado de la atracción generada por árboles matados con ácido cacodílico para manipular el comportamiento de los escarabajos de la corteza europeos e indígenas. Llevábamos dos años desarrollando este método del "árbol trampa", cuando abordamos los ensayos de campo en Syracuse, en 1978. A olmos muy enfermos o no deseados (considerados malas hierbas) situados en espacios verdes se les inyectó ácido cacodílico y se cebaron con la feromona. Los árboles sufrieron un fuerte ataque de ambas especies de escarabajo de la corteza, pero la reproducción del artrópodo falló de forma apreciable, si no por completo, debido a la sequía de la corteza interior inducida por el herbicida. Al final del tercer año (1980), la tasa de infección en el área de tratamiento había caído de 7,7 a 0,6 por ciento. En un área próxima utilizada como control experimental, la tasa de infección se mantuvo alrededor del 5 por ciento desde 1977 hasta 1979; cuando se extendió el tratamiento a este área, en 1980, la tasa de infección bajó al 2,2 por ciento.

En el futuro, a medida que se vaya controlando la enfermedad holandesa del olmo, el método del árbol trampa podrá abandonarse en favor de un saneamiento rápido y la caza masiva de escarabajos de la corteza. Al propio tiempo, podrán tomarse otras medidas de protección, siempre que la confrontación entre costes y expectativas de beneficios lo garantice. En armonía con la reducción del número de escarabajos de un área dada, los programas integrados de control de la peste necesitan nuevos procedimientos para aislar y dominar el patógeno.

La supervivencia de un árbol suele interesarle a quien lo tiene plantado en su jardín o a quienes aprecian la belleza de un paisaje. En los últimos años se han empleado fungicidas, basados en el

compuesto orgánico benzimidazol, para el tratamiento terapéutico y profiláctico de la enfermedad holandesa del olmo. El rendimiento se ha visto limitado por varias razones. Primero, inhiben el crecimiento de *Ceratocystis ulmi*, pero no lo matan. Segundo, ciertas razas del hongo oponen una resistencia total a tales fungicidas. Tercero, los fungicidas deben suministrarse repetidamente al árbol, porque su efecto se diluye a medida que éste crece. Se ha demostrado que esos tratamientos persistentes debilitan el árbol y permiten la entrada de otros patógenos a través de los lugares de tratamiento.




EL EFECTO ANTIFUNGICO de *Pseudomonas syringae* en olmos jóvenes se confirmó en este experimento. El árbol de la derecha se trató con un medio de cultivo que contenía una cepa bacteriana productora de un antimicótico. Al árbol control de la izquierda se le inyectó un medio de cultivo idéntico en el que no había crecido bacteria. (En ambos casos, el punto de inyección es la mancha negra de la parte inferior.) Dos semanas más tarde, se les aplicó a ambos árboles una dosis masiva de esporas y fragmentos de hifas de *Ceratocystis ulmi*. Un mes después se peló la corteza de los árboles. La madera del árbol control estaba completamente descolorida por la acción del hongo, mientras que, en el árbol tratado, las zonas situadas por encima y por debajo del punto de inyección aparecían libres de tal decoloración. De las zonas descoloridas de la madera se recuperó *P. syringae*. En el árbol tratado, el hongo probablemente se movió entre las zonas de la madera donde se había establecido la bacteria, dando a la madera un aspecto veteado.

En parte para evitar tales problemas, un grupo de investigadores dirigido por uno de nosotros (Strobel) comenzó a trabajar en el estado de Montana en un nuevo método biológico para tratar olmos aislados. Al observar que ciertas bacterias del género *Pseudomonas* sintetizaban antimicóticos, agentes antifúngicos, buscamos de entre una colección de estos organismos los que fueran antagonistas de *Ceratocystis ulmi*; encontramos varios. Puesto que algunos miembros de este género de bacterias se hallan frecuentemente asociados a vegetales, esperábamos que alguno de los organismos productores de antimicóticos se acomodaran en el árbol y le confirieran cierto grado de resistencia frente al *C. ulmi*. Seleccionamos entonces una raza de *Pseudomonas syringae*, muy eficaz como inhibidor de *C. ulmi*. Observamos que podía crecer en la savia del olmo y producía la sustancia antimicótica cuando se cultivaba sobre medio de agar que contenía savia; además, ni la bacteria ni el antimicótico semipurificado resultaban tóxicos para el olmo. Demostramos también que la bacteria se movía a veces dentro del árbol junto con el flujo regular de la savia y se establecía a lo largo de todo el árbol.

Administramos cultivos de la bacteria a plántulas de olmo de un invernadero, a mediados de mayo. Dos semanas después, se inocularon los arbolitos con un cultivo patógeno de *Ceratocystis ulmi*. Cuando se examinaron, ocho semanas más tarde, ninguno de los ejemplares tratados con la bacteria mostraba síntomas de la enfermedad holandesa del olmo. Los árboles control expuestos solamente al hongo o tratados con bacterias que no producían antimicóticos mostraban síntomas de la enfermedad. La eficacia de *Pseudomonas syringae* en el control de la enfermedad parece relacionarse con la capacidad de la bacteria para producir antimicóticos: en efecto, un mutante que no los sintetizaba se mostró completamente ineficaz a la hora de controlar la enfermedad.

Se confía también en el tratamiento biológico de árboles maduros enfermos del campo. En determinado experimento, se trataron con la bacteria 22 árboles enfermos, dejando, inermes, a modo de control, un número igual de árboles enfermos. Todos los árboles control, menos uno, murieron o se debilitaron rápidamente durante dos estaciones de crecimiento. En el grupo de los tratados, siete de los ocho árboles que estaban poco infectados y fueron tratados al principio de la estación de crecimiento sobrevivieron sin daño o con muy pocos síntomas de decadencia después de dos estaciones de crecimiento. Los otros árboles tratados, que habían recibido esos cuidados ya más avanzada la estación, o estaban fuertemente infectados, murieron o se degradaron. Concluimos provisionalmente que el tratamiento tenía más éxito cuando se efectuaba en época temprana, durante el tiempo de máximo flujo de savia en el árbol. Tres de los árboles tratados, aparentemente, se habían recuperado del todo, como probaba el hecho de que no se pudo volver a aislar *Ceratocystis ulmi* de ellos.

Hasta la fecha, miles de árboles de los Estados Unidos y Canadá se han tratado con *Pseudomona syringae*. Al refinar los métodos de tratamiento en futuros ensayos habrá que considerar varios factores; la capacidad de la bacteria para sobrevivir y difundirse en diferentes tipos de olmos; la remota posibilidad de que la cepa de bacterias utilizada sea patógena para ciertos vegetales y la posibilidad de que otras cepas de *Pseudomonas* resulten más efectivas, o que la cepa ensayada pueda mejorarse genéticamente. Además, deben calibrarse los posibles efectos de *P. syringae* sobre el hombre y el medio ambiente.

Se está ensayando en varias comunidades de los Estados Unidos un programa integrado de control de la plaga que comprende medidas de saneamiento de los árboles con fungicidas y bacterias antagonistas. Aunque las nuevas técnicas biológicas son prometedoras, se necesitan más ensayos. Mientras tanto, podemos estar seguros de que un considerable número de olmos se salvarán mediante diversos programas de control. Incluso en el campo, donde no se aplican cuidados, el olmo americano no está destinado a la extinción. Su tremendo poder de reproducción seguramente lo salvará. A la larga se puede esperar que la selección natural, o los programas de cruzamiento, permitan crear un olmo americano capaz de defenderse de *Ceratocystis ulmi*, como lo hacen sus parientes asiáticos.

Ciencia y Sociedad

¿Tartessos?

Las excavaciones de urgencia que el autor de esta nota (Juan Pedro Garrido) ha realizado en el solar n.º 10 de la calle del Puerto en la ciudad de Huelva han renovado, una vez más, la actualidad del tema de Tartessos. El hallazgo de interesantes niveles arqueológicos, correspondientes al siglo VI a. de C., con materiales y restos arquitectónicos que se relacionan cronológicamente con la rica necrópolis orientalizante en curso de excavación, ha suscitado el interés en los más diversos medios, planteando su conexión con el problema de Tartessos, que no puede eludirse por razones geográficas, cronológicas y arqueológicas.

Debemos tener en cuenta que el problema de Tartessos tuvo su divulgación fuera del marco de los especialistas desde las páginas literarias de la versión en castellano del "Tartessos" de A. Schulten, alcanzando una resonancia y popularidad nada frecuentes en los temas científicos. No obstante, existe una desorientación sobre este tema, provocada por el desenfoque del mismo, al haberse divulgado un planteamiento exclusivo o casi exclusivo de problema histórico en sentido estricto y no en los términos de arqueología prehistórica, que es el campo básico donde debe dilucidarse el problema.

Este planteamiento en términos de historia clásica (interpretación de fuentes escritas) se debe a que el tema de Tartessos fue abordado ya en el siglo XVI por el jesuita sevillano Juan de Pineda, en una época en que la arqueología no era fuente ni método histórico, sino sólo actividad de anticuarios y pasatiempo de curiosos y en la que no existía el propio concepto de la prehistoria como ampliación y superación de la historia. Por tanto, lo que hizo el iniciador del tratamiento del tema de Tartessos, Pineda, y con él todos los eruditos e historiadores posteriores, fue intentar reconstruir la historia en base a la documentación escrita. Ciertamente, este era el único método posible en los momentos en que ni la arqueología ni otras ciencias instrumentales de la historia se habían elaborado, ni aún el propio concepto de la historia había sido ampliado por la prehistoria y la etnología, disciplinas que no tomarán carta de naturaleza científica hasta fines del siglo XIX.

El problema básico de la aplicación de este método al tema de Tartessos radica en que los textos son escasos, oscuros y poco explícitos, y, por otra parte, proceden por lo general de fuentes indirectas y de épocas cronológicas posteriores. Los pocos datos que ofrecen las fuentes son, por las razones antedichas, totalmente insuficientes para reconstruir los aspectos históricos de Tartessos. No obstante, aún en épocas recientes y pese al avance y nuevas orientaciones de las ciencias históricas, no faltan intentos de determinar el significado de Tartessos, su localización geográfica y no pocos hechos históricos, en base a la interpretación de los aludidos textos. Con ello no pretendemos negar la utilidad de las fuentes escritas, sino sólo manifestar nuestro escepticismo cuando se utilizan como única y exclusiva fuente de conocimiento histórico, o, de forma apriorística, tratando de acomodar o interpretar los textos en conformidad con ideas y conceptos previamente elaborados. De esta manera se ha considerado dogmáticamente que la Tharshish citada en los textos bíblicos equivale a la Tartessos de las fuentes griegas (lo que en todo caso es un problema discutible), complementando la exégesis con los escasos y oscuros o imprecisos textos griegos; de esta forma se ha localizado Tartessos en la Syrtes, en el Mar Rojo, etcétera, y, dentro de la península ibérica, en el extenso espacio comprendido entre Cartagena y Huelva.

Incluso existe un texto sobrevalorado en exceso desde nuestro punto de vista. Me refiero al periplo de Avieno, obra del siglo IV d. de C., que parece basarse en un antiguo rotero massaliota del siglo VI a. de C., que llegó a manos del poeta Rufus Festus Avienus a través de refundiciones y reelaboraciones posteriores del texto primitivo. Desde que el filólogo A. Schulten intentó románticamente emular a Schliemann descubriendo una nueva cultura porque así podía deducirse de las fuentes literarias, han sido muchos los que han reelaborado tesis en base a la interpretación de dichos textos, ya sea para rebatir, confirmar o modificar las ideas de Schulten. Ciertamente, Schulten buscó la confirmación arqueológica de su tesis sobre la localización de Tartessos en el Coto de Doñana, por lo que en compañía de Jorge Bonsor realizó excavaciones en el lugar donde teóricamente, en su opinión, debía encontrarse Tartessos, y en el que, según la memoria redactada por Bonsor, sólo aparecieron restos romanos tardíos y un solo elemento que podía ser de época tartéssica: un anillo con inscripción griega arcaica.

Después de estas excavaciones de Bonsor y Schulten, la bibliografía dedicada a la interpretación de las fuentes escritas, adecuándola con la geografía actual y determinado el lugar identificable con el desaparecido Tartessos, ha sido excesiva. Pero desde hace algunos años, debido en parte al escepticismo sobre el valor de las fuentes escritas para resolver el problema y, en parte, a espectaculares hallazgos casuales acaecidos en Andalucía, se ha centrado el problema de Tartessos en términos estrictamente arqueológicos, dentro del marco de la prehistoria. A esta nueva orientación contribuyeron los estudios de especialistas, que procedieron a valorar una serie de elementos materiales ya conocidos, dejando a un lado los textos escritos a los que sólo se les concedía un valor relativo y supletorio. Así, los trabajos de Almagro sobre el depósito de bronces hallado en la ría de Huelva en el año 1923 y la valoración cronológica y cultural de la fíbula de codo tipo Huelva, junto con los estudios sobre jarros de bronce y otros materiales orientalizantes de Blanco Freijeiro, García y Bellido, Cuadrado y Blázquez marcan importantes hitos en el planteamiento del tema de Tartessos desde perspectivas estrictamente arqueológicas.

En la actualidad, nuestra investigación se orienta en el sentido de buscar el dato material mediante la excavación arqueológica, aunque sin desechar por completo el valor orientativo general que ofrecen los textos escritos, y, sobre todo, en los aspectos deducibles de una interpretación basada en métodos etnológicos, como hace algún tiempo realizado J. Caro Baroja, que si bien no responden al interrogante de la localización topográfica de Tartessos, sí nos informan sobre su significado cultural y alto grado de civilización.

Por ello han sido del mayor interés los hallazgos de la calle del Puerto, aunque no hayan constituido sorpresa sino algo previsible y esperado por simple deducción lógica a partir de los datos obtenidos en las excavaciones metódicas de los cabezos de Huelva (San Pedro y La Esperanza) y en la rica necrópolis del período orientalizante (siglos VII-VI a. de C.). Además, la recuperación de significativas cerámicas y otros elementos arqueológicos arrojados a los vertederos de escombros proceden-

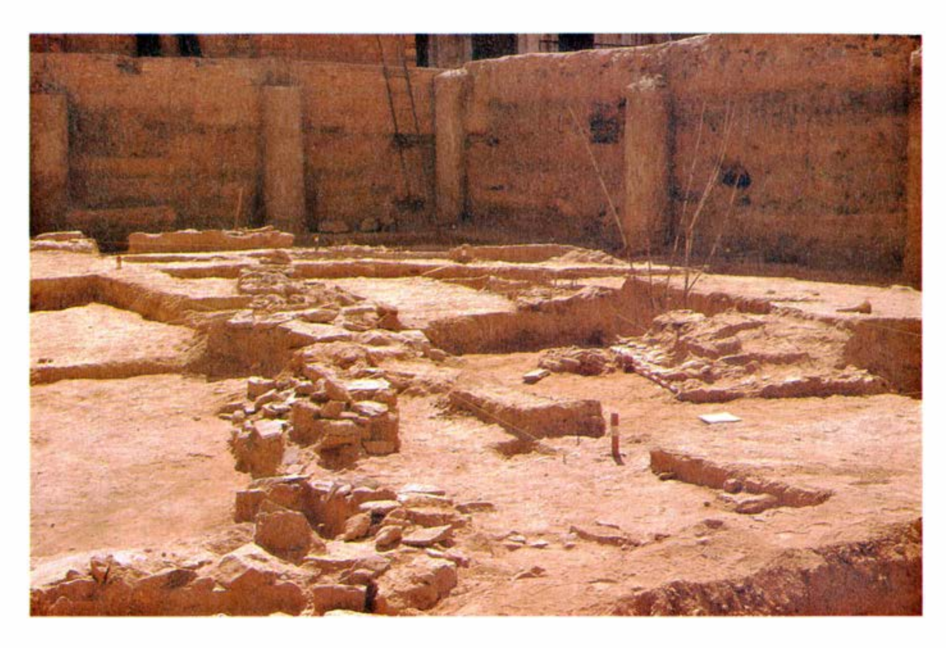
tes de las obras de edificación de las calles Méndez Núñez y Rafael López, aparte de otros hallazgos, permitían reforzar nuestro punto de vista, compartido por otros arquéologos, de que en el subsuelo de la parte antigua de Huelva debían encontrarse los restos de un hábitat que pudiera correlacionarse cronológicamente con la excepcional necrópolis orientalizante.

La novedad la ha constituido el hallazgo de una superposición de niveles de habitación en una estratificación de más de seis metros, que documentan la existencia de un denso hábitat con continuidad desde los siglos VIII o IX a. de C. hasta época romana, donde se conservaban restos arquitectónicos y otros elementos de excepcional valor científico.

Lamentablemente, nuestra intervención tuvo lugar con posterioridad a la destrucción de los niveles más recientes (siglo V a. de C. a época romana) por parte de la empresa constructora, que había iniciado los trabajos de cimentación en el solar correspondiente al número 10 de la onubense calle del Puerto.

Nos personamos el día 18 de agosto de 1980 en el solar de referencia, donde contemplamos un rectángulo de más de 400 metros cuadrados, del que la empresa constructora había extraído tierras en su totalidad hasta una profundidad de 3,20 metros y perforado el suelo en varios puntos donde se habían colocado pivotes de sustentación. En las paredes resultantes se apreciaba con toda nitidez la existencia de estratos arqueológicos y muros de mampostería, que en la parte occidental quedaban bajo una potente capa de hormigón romano sobre el que se habían asentado los cimientos de las casas medievales y de comienzos del siglo XX. En consecuencia, hubimos de suspender la excavación de la necrópolis donde habíamos iniciado los trabajos en dos nuevas sepulturas tumulares, a fin de trasladar al equipo técnico al solar de la calle del Puerto para documentar los hallazgos. Allí, y a partir de la profundidad citada de 3,20 metros, iniciamos nuestros trabajos cuyos resultados habrían de superar los cálculos más optimistas: inmediatamente, se encontró un nivel arqueológico constituido por una serie de abigarradas y complejas construcciones que fue posible fechar en el siglo VI a. de C. por la presencia de ricas y abundantes cerámicas.

Estas construcciones pueden identificarse con almacenes, con toda probabilidad en conexión con el comercio marítimo, pues las orillas del estuario debían estar muy próximas en esa época. En este mismo nivel y en el ángulo noroeste de la excavación, se encontró parte de un edificio de características diferentes, cuyas paredes presentaban un aparejo distinto y al que probablemente habrá que atribuir distinta funcionalidad. ¿Quizás un templo? Parte de esta gran construcción continúa por debajo del solar colindante, lugar que nos proponemos excavar tan pronto se cuente con medios para ello. Hemos de señalar que en este nivel hay ricas cerámicas griegas que indican la potencia económica de las gentes que aquí vivieron. Entre éstas destaca un fragmento que el especialista del Museo Arqueológico Nacional, Olmos, atribuye al pintor Clitias, el autor miniaturista que decoró el vaso François, y fechable, por consiguiente, hacia el 575 a. de C. El taller alfarero que lo elaboró debió de ser el de Ergotimos. Se trata de un fragmento correspondiente a un olpe en cuya superficie se representó la escena del juicio de Paris. En el fragmento de Huelva vemos parcialmente a la diosa Atenea, que ha de ir precedida probablemente por la diosa Hera, y el nombre que identifica a la diosa, en caracteres griegos, dispuesto verticalmente con las letras orientadas de derecha a izquierda. El significado y valor de este hallazgo es múltiple: por un lado ofrece una cronología muy definida; por otro, tiene un valor artístico intrínseco; finalmente denota la aludida riqueza y po-



Nivel en el que se inició la investigación en la calle del Puerto. En las paredes se aprecia la sucesión de estratos arqueológicos bajo una capa de hormigón romano

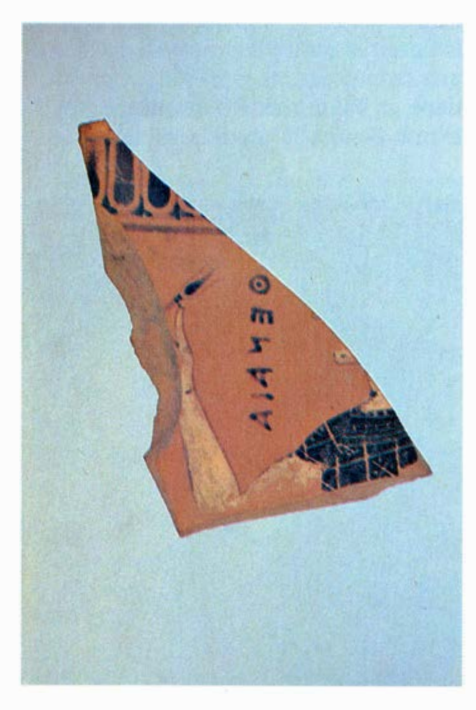


Construcciones complejas del primer nivel investigado correspondiente al siglo VI antes de Jesucristo



Edificio del ángulo noroeste cuyos muros continúan bajo el solar colindante por el W

tencialidad económica que permitía una importación excepcional, valiosa y costosa. Este punto de vista queda reforzado por el hallazgo de fragmentos de cerámicas griegas de la misma época, de buena calidad y en relativa abundancia, también atribuibles a otros pintores griegos de nombre conocido, y que hacen pensar en la posibilidad de unos contactos directos con el mundo griego en esta época, probablemente con los focenses, sin mediación de los fenicios, tal como nos han transmitido las fuentes escritas. En todo caso, constituyen el más importante conjunto de cerámicas griegas antiguas halladas en el Sur de España. Por debajo de este nivel y, por tanto, de fecha más antigua, se apreció la existencia de otras construcciones mejor conservadas y de aspecto más sólido. En este caso parece tratarse de viviendas de planta rectangular y paredes de pizarra. Los pisos, en los que se encuentra algún hogar, están teñidos de rojo, probablemente con almagre. Tanto las paredes de pizarra como el teñido de rojo de los suelos son, todavía hoy, elementos característicos de la arquitectura popular en la próxima región de El Andévalo. También en este mismo nivel abundaban las escorias de fundición de mineral de plata, y hasta apareció un horno de fundición, detectándose la presencia de otros dos que permanecen sin excavar.



Fragmentos de un olpe pintado por Clitias, el más famoso artista cerámico ateniense de la época arcaica. Se trata de una pieza excepcional y única en la península ibérica

Pero más sorprendente fue comprobar que, todavía por debajo, aparecía un tercer nivel de construcciones que no hubo oportunidad de excavar en extensión. Parece que los estratos de este nivel van asociados a una variedad cerámica llamada de "retícula bruñida", que comúnmente se considera indígena (o no fenicia) y cuya cronología se remonta al siglo IX a. de C.

Debemos señalar que estos tres niveles parecen presentar unas características de urbanización de calles trazadas a cordel, con una concepción urbanística propia de las altas culturas del Mediterráneo oriental, debida lógicamente, al igual que otros elementos arqueológicos, a la presencia o influencia de gentes de este origen.

La vinculación de los primeros niveles arqueológicos que nos ha sido posible investigar en la calle del Puerto con la necrópolis que estamos excavando en los cabezos de Huelva nos ofrece un amplia perspectiva sobre esta época en la que las fuentes escritas sitúan el apogeo de Tartessos.

La ingente cantidad de objetos depositados como ajuar funerario en las tumbas evidencia el alto grado de refinamiento y civilización alcanzado en los comienzos del siglo VI a. de C.: joyas de oro, marfiles y alabastros, junto con depurados objetos de bronce de perfecta ejecución técnica y artística nos revelan el elevado grado de cultura y la potencialidad económica alcanzada. Ahora, con el hallazgo de la potente estratificación encontrada en la calle del Puerto, que constituye sólo una muestra de lo existente en el subsuelo del casco antiguo de la ciudad, tenemos la prueba material de la existencia de un hábitat urbano plenamente desarrollado en el siglo VI a. de C., cuyos orígenes se remontan, al menos, a dos siglos antes. La relativamente gran cantidad de valiosas cerámicas griegas parecen confirmar la tradición literaria sobre el filohelenismo de Argantonio, el rey de la plata; las escorias de mineral de plata halladas aquí y en otros lugares de la ciudad apuntan igualmente al fondo de realidad que puede esconderse en los textos escritos.

En consecuencia, en el estado actual de la investigación, podemos adecuar los datos arqueológicos y geográficos con los testimonios literarios sobre Tartessos, para establecer las siguientes conclusiones:

1) Queda confirmado, desde un punto de vista arqueológico, la existencia de un denso y extenso hábitat, donde, y a partir de fecha temprana, observamos un notable desarrollo del urbanismo. Existe una concentración urbana que podemos considerar una gran ciudad, pues su perímetro es superior al de Huelva a comienzos de siglo. (Las fuentes escritas nos hablan de Tartessos como de una ciudad.)

2) En dicho hábitat abundan los indicios que nos indican la existencia de actividades metalúrgicas encaminadas a la obtención de plata. (Las fuentes escritas nos hablan de la riqueza en plata de Tartessos. El nombre de Argantonio parece equivaler a rey de la plata.)

3) La variedad, calidad y diversa tipología de las cerámicas encontradas en el hábitat, con ánforas y vasos de gran tamaño, indican amplias relaciones comerciales; lo mismo cabe decir con respecto a las joyas y a otros elementos de inspiración oriental que encontramos en los ajuares funerarios de la necrópolis. La plata y otros metales obtenidos en las fundiciones habrían de ser comercializados. (Las fuentes escri-

tas nos hablan de Tartessos como de un mercado.)

4) La ciudad y el mercado están emplazados al pie de una montaña entre los dos brazos del doble estuario del Tinto-Odiel, hoy notoriamente reducido con respecto a la mayor amplitud en otras épocas. (Las fuentes escritas nos hablan de un río y de la ciudad que está entre los dos brazos como en una isla. Debemos tener en cuenta que no hay vocablo equivalente a península en lengua griega; por tanto, lo mismo puede traducirse el término griego νησος por isla como por península. Esto conviene particularmente al yacimiento arqueológico de Huelva, situada en la parte meridional de una península que rodean no sólo los ríos Tinto y Odiel, sino también el Anicoba, dejando un reducido y estrecho istmo que fue hasta el siglo pasado la única vía de acceso terrestre.)

5) Los amplios estuarios que circundan la ciudad son las vías naturales de penetración hacia las muy cercanas fuentes metalíferas. Ambos ríos arrastran en sentido real sales metalíferas en disolución, como se ha hecho notar por varios investigadores. Pero también en un sentido no menos real debían descender los metales por vía fluvial, ya que hasta épocas relativamente próximas los ríos eran navegables en mayor trecho que en la actualidad, para confluir en la península de Huelva. (Las fuentes escritas nos hablan de que el río Tartessos arrastraba metales. No parece ser una metáfora poética.)

6) Las únicas minas de la península ibérica que en la antigüedad ofrecían a la vez, y en grandes cantidades, oro, plata y cobre fueron las de la región de Huelva. Según I. Pinedo, el análisis comparativo de las masas de mineral (proporcional en este caso a las de los escoriales antiguos) permite cubicar la existencia de mineral limpio en la región de Huelva por "centenares de millones". En comparación con la zona minera de Huelva, las minas de Sierra Morena (cobre de Peñaflor y Cerro Muriano, plomo y plata de Linares-Carolina) representan muy poco; para la comparación hay que recurrir a la escala de la pulga y el elefante. Fue tan intensa en Huelva la explotación minera en épocas remotas que los escoriales antiguos se están reexplotando hoy, dado el elevado porcentaje de metales preciosos recuperable por toneladas métricas. (Las fuentes escritas nos dicen que los metales tartéssicos eran la plata, el cobre y el oro.)

7) Entre los objetos hallados en la necrópolis orientalizante existen algunos fabricados en latón, es decir, de cobre y zinc en lugar de estaño. El zinc como metal independiente no se ha conocido hasta el siglo XVIII de nuestra era. Los artífices de estas piezas usaron intencionadamente minerales con fuerte contenido de zinc para manufacturarlas, obteniendo así una apariencia áurea. Las mayores reservas de zinc se encuentran en los llamados "minerales complejos" de la región de Huelva, donde se reúnen las tres cuartas partes de las reservas totales de zinc; como ha puesto de manifiesto I. Pinedo, "en una pequeña extensión de kilómetros cuadrados del Suroeste hay doble cantidad de plomo y triple de zinc que en el resto del país"; todo ello a muy poca distancia de la ciudad-mercado y del río. (Las fuentes escritas nos hablan de un bronce de características especiales, el "tar-

tessiós chalkós", el bronce tartéssico. Creemos que bien pudiera tratarse de latón obtenido en este área.)

8) Es evidente la unidad cultural y relaciones en el siglo VI a. de C. con una amplia zona del interior y del Valle del Guadalquivir. Ignoramos si ello implica una unidad política. No obstante, debemos tener en cuenta que una sociedad entre cuyas componentes figura la pastoril necesita un cierto territorio sobre el que desarrollar esta actividad económica, ejerciendo una cierta soberanía "de facto". (Tanto la importancia de la ganadería como lo noción de imperio o dominio territorial están documentadas por las fuentes escritas.)

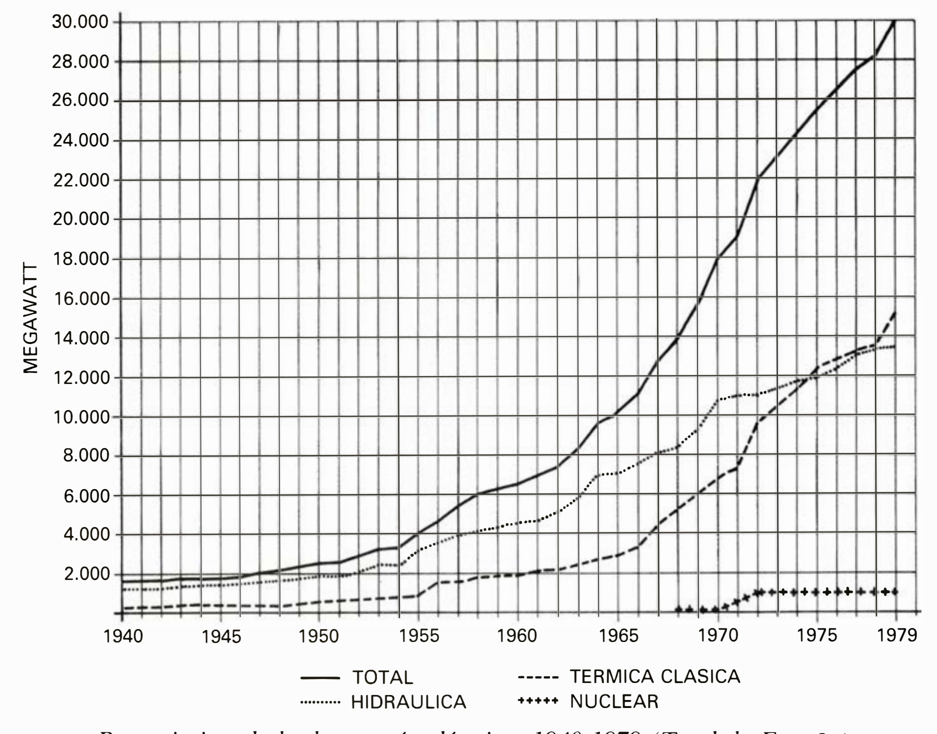
Vemos, pues, en razón de lo expuesto, que los vestigios arqueológicos de Huelva, contemplados desde la triple perspectiva cronológica, geográfica y arqueológica, parecen dar contenido material a lo que las fuentes escritas nos han transmitido con el nombre de Tartessos. Podemos precisar que esta cultura tartéssica no era resultado de la evolución de una cultura indígena que se hubiera desarrollado por su propia dinámica. El estímulo que había provocado el florecimiento del urbanismo y la acumulación de documentos arqueológicos era de origen oriental, entre los que los fenicios desempeñaron, al parecer, un papel fundamental, aunque no único ni exclusivo. El motivo de la atracción de elementos orientales parece ser la explotación y comercio de los metales, según refieren las fuentes escritas. Si nos atenemos estrictamente a la valoración de la cultura material, la componente indígena nos aparece un tanto diluida. No obstante, no podemos desconocer la componente pastoril, que puede representar la aportación indígena a este complejo cultural. Las culturas pastoriles no suelen ser muy expresivas en cuanto a restos materiales, pero pueden haber dado su personalidad y carácter a lo "tartéssico". Algo así puede deducirse sobre la jerarquización de la sociedad que reflejan los enterramientos de la necrópolis. Esperamos que las investigaciones futuras sobre bases arqueológicas permitan precisar y aclarar, aún más, los todavía numerosos puntos oscuros del "problema tartéssico". (J. P. Garrido Roiz.)

La hidroelectricidad en España

Cuando en un país se ha producido una transformación de sus modelos económicos tradicionales y de forma paralela una elevación del nivel de vida humano, las necesidades energéticas han aumentado considerablemente; existe una correlación estrecha entre desarrollo y consumo de energía. Sin embargo, con frecuencia se asiste, a nivel de nación, a un auténtico desequilibrio entre producción y consumo, dando lugar a una dependencia externa, que en ocasiones ha tenido graves repercusiones socioeconómicas. La idea de lograr un autoabastecimiento en fuentes de energía es una constante mundial, y se manifiesta en el desarrollo de una investigación que persigue el "conocimiento" preciso para transformar energía ya existente pero hoy apenas utilizada, o el descenso de los costes de producción que también constituyen obstáculos graves. Más que de una "crisis energética" cabría hablar de una "crisis tecnológica o económica en materia energética".

Dentro de este subsector, la electricidad se sitúa en un plano de excepción, precisamente no sólo porque sus múltiples posibilidades de uso han originado un amplio y heterogéneo mercado, sino porque, dada su naturaleza, puede obtenerse por procedimientos diversos. La técnica posibilita la movilización y transformación en electricidad de una energía ya existente, siendo por ello muy amplio el abanico de bases que posibilitan su producción, si bien en el momento actual no todas se conocen o se sitúan en condiciones rentables de explotación. Uno de los procedimientos tradicionales para su obtención ha estado unido a la modificación de la energía mecánica desarrollada por un volumen de agua al caer; la llamada hidroelectricidad. Sus exigencias son mínimas: agua, recurso autóctono y renovable, junto a desniveles topográficos, que facilitan la construcción de instalaciones.

España cuenta en cierto modo con un medio natural idóneo para la producción de energía eléctrica de origen hidráulico, sobre todo, dadas las características geomorfológicas, climáticas e hidrológicas de su mitad Norte. Por ello, por ser una forma de energía no contaminante, porque ofrece una producción instantánea, y porque su proceso productor no envejece de forma inmediata, se puso en marcha su explotación. Iniciada a finales del siglo XIX ha tenido una trayectoria compleja hasta nuestros días, que no se define precisamente por su homogeneidad. Podríamos afirmar que tuvo un pasado glorioso, cuenta con un presente mediocre y estamos ante una incógnita sobre qué será su futuro. Tomando como punto de partida el año 1940, por ser el momento en que la producción va a experimentar un mayor despegue, se identifican una serie de etapas muy significativas. La primera iría de 1940 a 1957, y se define por un gran predominio en la estructura de la producción eléctrica nacional; en todos los años la potencia instalada por origen hidráulico es superior al 70 por ciento del total instalado. De 1958 a 1965 se mantiene la prioridad hidroeléctrica, si bien se observan ya algunos años en los que desciende su nivel de participación; empieza a notarse un mayor despegue de la térmica convencional. De 1966 a 1972 podríamos hablar de un equilibrio hidráulico-térmico y, a partir de 1973, de una pre-

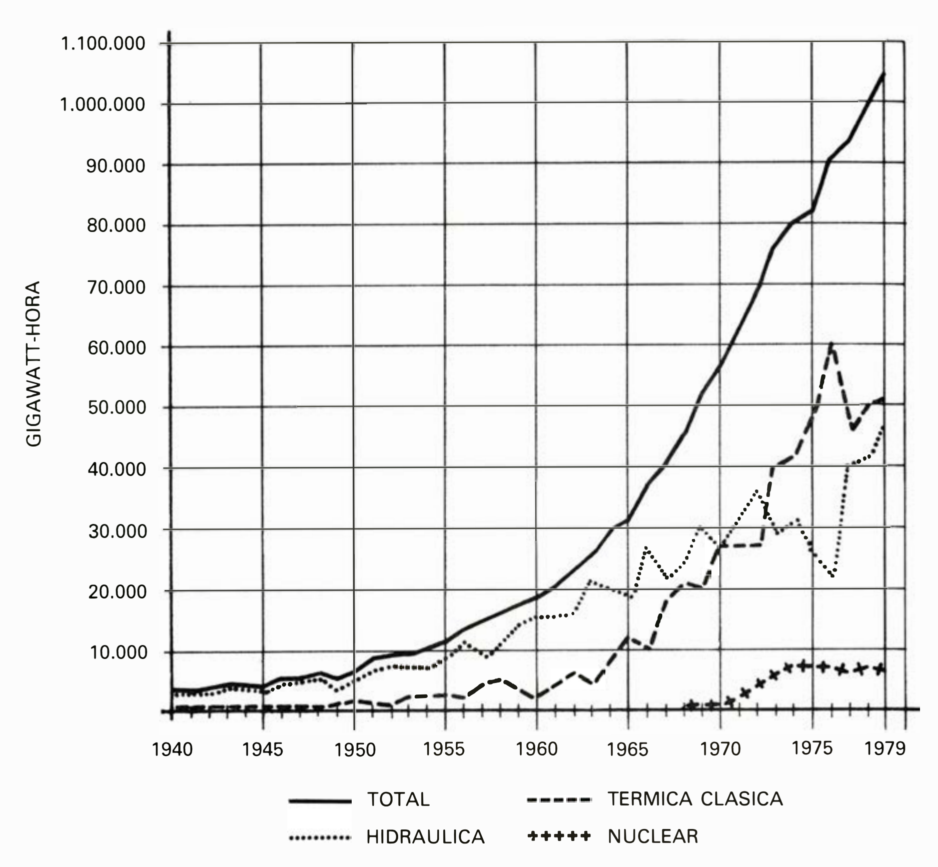


Potencia instalada de energía eléctrica, 1940-1979 (Total de España)

ponderancia termoeléctrica; la hidroelectricidad ocupa un lugar secundario. Se ha pasado de ser un país que basaba su producción eléctrica en el agua, teniendo a las centrales térmicas como complementarias, a poseer una estructura apoyada en la termoelectricidad, quedando lo hidráulico para los momentos de mayor demanda, lo que en la curva de carga se denominan horas punta. Es cierto que la potencia hidroeléctrica ha sufrido un incremento considerable; no en vano, de los 1340 megawatt de 1940 se ha pasado a los 4600 de 1960 y a los 13.580 de 1979; sin embargo, en la década de los años setenta se ha paralizado la construcción de centrales. Son varias las causas que explican esta, llamémosle, "decadencia" hidroeléctrica en España.

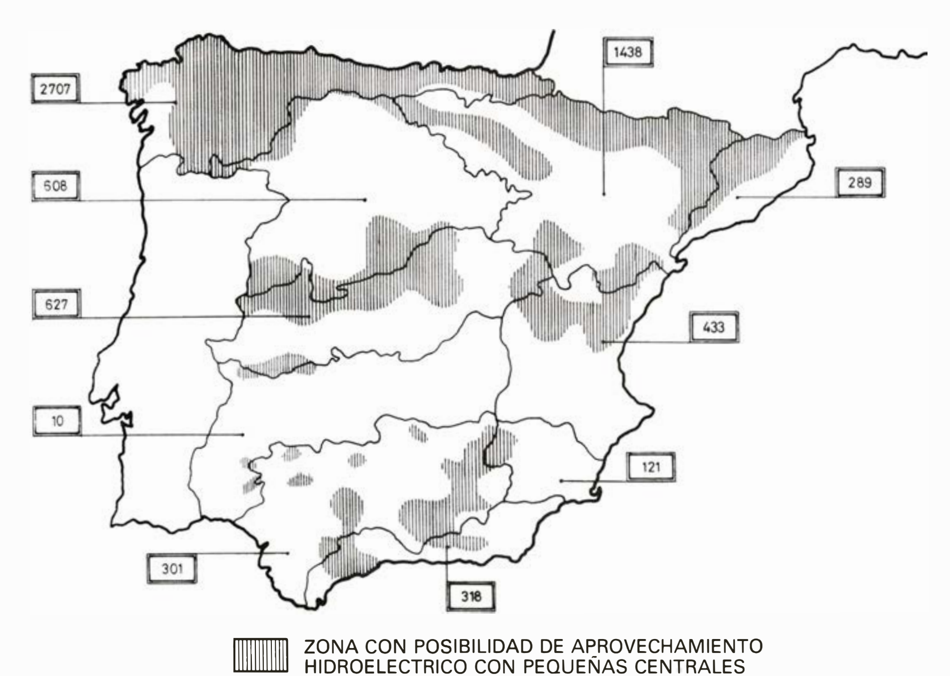
La demanda de energía eléctrica ha representado un papel trascendental. Su crecimiento ha sido importante; de los 22.799 gigawatt/hora de 1963 se ha pasado a 99.381 gigawatt/hora en 1979. Este fuerte consumo exigía un suministro constante y asegurado, aspectos que no cumple la hidroelectricidad, que se caracteriza precisamente por sus grandes fluctuaciones, ya que depende sensiblemente de la hidraulicidad del año. España, a pesar de que cuenta con zonas que acusan un matiz climático oceánico o continentalizado, se entronca dentro del contexto general mediterráneo, por lo que la mayor parte de sus ríos acusan una fuerte irregularidad, circunstancia que condiciona sensiblemente este tipo de producción. Por ello, era preciso desarrollar aquellos procedimientos que permitiesen un abastecimiento sin problemas; de ahí el impulso térmico, tanto convencional como nuclear. Así, la política energética nacional se orientó hacia la termoelectricidad, iniciando un cierto "proteccionismo" que benefició la construcción de este tipo de centrales. Las primas al combustible han conllevado a una valoración no real del precio del kilowatt/hora producido; los conciertos entre empresas productoras-Estado no han favorecido la utilización hidráulica. Ejemplo de ello sería el Decreto de Acción Concertada de 1975 con vigencia hasta 1985, que, al no contemplar los gastos reales por la adquisición de terrenos, desplazaba a las construcciones hidroeléctricas, ya que si van unidas a sistemas de regulación fluvial precisan de amplios espacios.

La propia Ley de Aguas ha constituido un freno importante al no considerar los usos integrales del recurso a que hace referencia, estableciendo un orden preferencial, en el cual la hidroelectricidad ocupa un lugar muy secundario. Este aspecto es importante, ya que, si se llevara a cabo, se lograría una mayor eficacia en la utilización del agua y, lógicamente, un mejor aprovechamiento de la infraestructura fluvial, de cara al desarrollo no sólo del sector primario (regadíos), sino también del secundario (energía). Una ley que data de 1879 difícilmente podrá adaptarse a las necesidades actuales.



Producción anual de energía eléctrica, 1940-1979 (Total de España)

A pesar de estos problemas, España ofrece una importante potencia instalada de origen hidráulico, localizada preferentemente en las cuencas del Norte, Ebro, Duero y Tajo. Ellas coinciden con ríos de alimentación pluvial oceánica o nival, más o menos matizada, lo cual les confiere un mayor caudal y regularidad. La mitad Sur del país y la vertiente mediterránea, a excepción de



Potencial hidroeléctrico español aprovechable con pequeñas centrales (Gigawatt-hora/año)

la cuenca del Ebro, se identifica con ríos de clara alimentación pluvial mediterránea, que se traduce en un caudal escaso e irregular; de ahí que se caracterice por un vacío productor. La utilización de los recursos hidráulicos españoles no ha llegado a su techo, todavía quedan posibilidades de cara al futuro, y es de prever que, si se modifican esos aspectos considerados como negativos para su desarrollo, la hidroelectricidad tenga de nuevo un mayor peso. Nunca será una energía de base, pero si contribuye en mayor medida junto con otras formas de producción a satisfacer la demanda energética irá disminuyendo la dependencia de otras fuentes, como el petróleo, que tanto está condicionando nuestra trayectoria económica.

Los recursos hidroeléctricos son todavía importantes. Pese a las dificultades que entrañan su evaluación, un último estudio realizado por la Dirección General de Obras Hidráulicas y UNESA demuestran su significación. El potencial hidroeléctrico bruto actual se cifra en 137.922 gigawatt/hora; el futuro, detrayendo el agua necesaria para otros usos, se cifra en 127.935 gigawatt/hora. Ahora bien, no todo él es técnicamente desarrollable, ya que se tropieza con problemas derivados del propio proceso productor, de las oscilaciones en el nivel de los embalses, de aquellos ligados al contexto económico (inundación de espacios habitados, económicamente rentables o de interés ecológico), de la geología o geomorfología de la zona o de la necesidad de dejar un mínimo de caudal fluvial. Por ello, teniendo en cuenta estos aspectos, ese potencial disminuye sensiblemente, quedando en unos 27.205 gigawatt/hora, como nuevos recursos, lo cual supone aproximadamente un 40 por ciento más de lo que actualmente se produce. Las zonas donde se localizan coinciden de nuevo con las cuencas del Norte, Ebro, Duero y Tajo.

Por otra parte, existe la posibilidad de explotar los mínimos caudales, mediante las llamadas pequeñas centrales automáticas. Así, la Dirección General de Obras Hidráulicas, la Comisaría de Energía y UNESA han elaborado un estudio al objeto de constatar esa utilización. Se trata de "minicentrales" instaladas directamente sobre el caudal fluvial, son saltos fluyentes, se adaptan al volumen de agua disponible y exigen escasos caudales, equivalentes como máximo a 6,35 metros cúbicos por segundo, si bien es un valor susceptible de ampliación. Cuentan con una potencia que oscila entre los 250 y 5000 kilowatt y su utilización se sitúa en unas 5000 horas al año. Permiten aprovechar esas pequeñas corrientes fluviales ligadas generalmente a los sistemas montañosos, que en el momento actual habían quedado desestimadas. Los recursos evaluados se sitúan en 6852 gigawatt/hora al año, siendo las corrientes norteñas, pirenaicas y las del Sistema Central las que ofrecen mayores posibilidades. Ya se ha iniciado un plan de actuación por las sociedades eléctricas que integran UNESA, y está prevista la construcción de 113 "minicentrales" con una inversión inicial de 3600 millones de pesetas, si bien se estima que el total ascenderá a unos 6500 millones.

Todavía existen en España recursos hidroeléctricos que deben ser utilizados. La política energética nacional ha de tenerlos en cuenta y dedicar su atención a ellos, ya que a pesar de que su desarrollo no lograra satisfacer todas nuestras necesidades energéticas, sí podría paliar en cierto modo la crisis actual. El Plan Energético Nacional dedica una inversión pequeña para el nuevo impulso hidroeléctrico; está previsto que en 1983 tan sólo el 8 por ciento del total invertido se dedique a esta forma de obtención de electricidad; quizás ante los nuevos conocimientos pudiera modificarse esta cifra de forma positiva. La hidroelectricidad puede y debe contribuir al incremento de la producción de energía eléctrica; es preciso utilizar un recurso propio que además indirectamente transformaría las condiciones socioeconómicas de las áreas afectadas. Generalmente, los centros productores de electricidad, en España, acusan un importante subdesarrollo, sin que esa energía haya supuesto beneficio alguno, pese a tener que soportar las servidumbres de una central. La Ley del Canon aprobada el 10 de marzo de 1981, trata de compensar económicamente a esos centros de producción, gravando sobre el consumidor un 5 por ciento del precio medio del kilowatt/hora. La cantidad distribuida a las diputaciones provinciales en razón de la potencia instalada de su demarcación deberá emplearse en mejorar la infraestructura de las áreas productoras. Si la gestión es eficaz, podrá repercutir beneficiosamente en la ordenación de algunos espacios, por ejemplo zonas de montaña, que precisan una transformación urgente de su estructura económica, siendo importante para ello la génesis de unas bases físicas adecuadas (la propia infraestructura citada). La utilización del agua puede aminorar el problema energético actual, y también actuar como motor de desarrollo de tantas zonas actualmente deprimidas tanto económica como demográficamente. (Mercedes Molina Ibáñez.)

El ribosoma

A partir de electromicrografías se ha diseñado un modelo en tres dimensiones de la máquina intracelular que sintetiza proteínas. Se contemplan las secuencias necesarias para producir una proteína

James A. Lake

La función que desempeña en la célula el minúsculo gránulo llamado ribosoma es fácil de resumir: recibe instrucciones genéticas y las traduce a proteína. En la primera parte de esta tarea, procesa la información. Se une al ARN mensajero, una molécula lineal cuya secuencia de unidades, llamadas bases nucleotídicas, constituye un código para los aminoácidos. A continuación, incorpora los aminoácidos en la secuencia adecuada. La segunda parte de esta tarea es esencialmente catalizadora: une los aminoácidos en una cadena para formar una proteína.

En un primer sentido podemos decir que el ribosoma es pequeño. Su dimensión mayor es de 250 unidades angstrom, 25 millonésimas de milímetro, aproximadamente, lo que supone alrededor de la décima parte del tamaño de una longitud de onda de luz visible. Por tanto, los ribosomas son invisibles al microscopio óptico, excepto como lunares en las células que contienen gran número de ribosomas y polisomas. Los polisomas no son otra cosa que un conjunto de ribosomas agrupados para traducir una cadena sencilla de ARN mensajero.

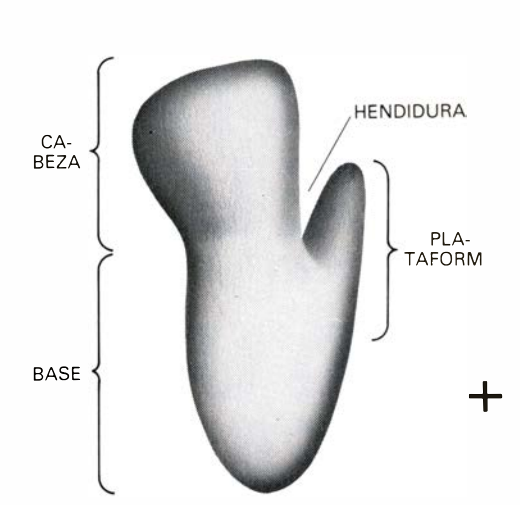
En otro sentido el ribosoma es grande. Una de las principales técnicas que se emplean para deducir la estructura de una molécula es la difracción de rayos X, en la que la repetición de la disposición tridimensional de los átomos en un cristal de una sustancia dada origina un patrón de difracción, o patrón de interferencia, cuando un haz de rayos X se dirige a través del cristal. El ribosoma, sin embargo, tiene el peso de 100 a 150 moléculas de proteína de tamaño medio. Constituye una máquina formidable; la mayoría de los investigadores lo consideran un orgánulo intracelular. Para una estructura tan grande (y compleja) como la del ribosoma, el desarrollo de un cristal para difracción de rayos X se ha supuesto, hasta hace poco, tarea imposible. El patrón de difracción producido por el cristal resultaría de extremada complejidad.

Y, en un tercer sentido, el tamaño del ribosoma es ideal. El microscopio electrónico puede resolver ya dos puntos de una estructura que estén separados unos 15 angstrom. Con esta resolución, una molécula de proteína de 30 angstrom de diámetro aparece más o menos esférica o elíptica en las electromicrografías. Con la misma resolución, un ribosoma nos proporciona una imagen compleja.

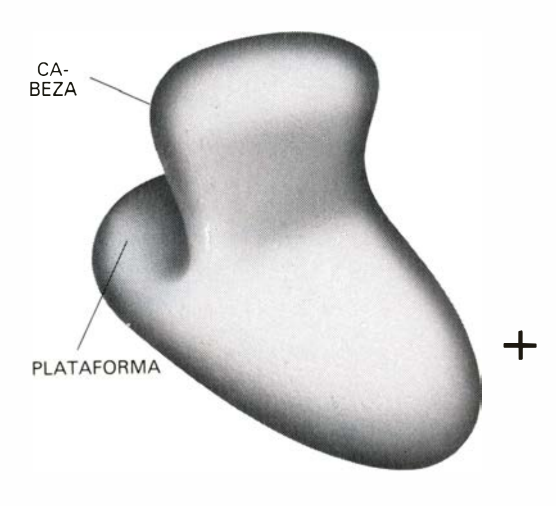
El tamaño del ribosoma supone que, para dilucidar su estructura tridimensional, había que ordenar varios miles de datos diferentes. Ni siquiera un mapa que señalara la localización de cada átomo en el ribosoma revelaría, por sí mismo, la forma en que el ribosoma fabrica proteínas. Las técnicas que describiré permiten la identificación de lugares funcionales en el ribosoma. Los descubrimientos sobre la función del ribosoma pueden "cartografiarse", pues, en su estructura. Se espera, desde luego, que la elucidación de la estructura, junto con el descubrimiento de los sitios funcionales, sugiera, hasta en sus pormenores, el mecanismo molecular según el cual trabaja el ribosoma. Al final de este artículo señalaré la ubicación de varios sitios funcionales y propondré las líneas generales de un posible mecanismo. El primer problema, sin embargo, era averiguar la morfología del ribosoma.

Descubrimiento del ribosoma

Ya que el tamaño del ribosoma permite su observación al microscopio electrónico, en éste han hallado los biólogos una fuente importante de información sobre la estructura del orgánulo, a lo largo de las tres décadas transcurridas desde que se descubrió el ribosoma. A decir verdad, tal hallazgo fue posible merced al microscopio electrónico. A principios de la década de 1950, Albert Claude, Keith R. Porter y George Palade, del Instituto Rockefeller de Investigaciones Médicas, observaron al microscopio electrónico que, en el citoplasma de las células, abundaban los gránulos esféricos. Una bacteria típica contenía unos 10.000 gránulos; las células de los organismos superiores encerraban muchos más. Los ha-



SUBUNIDAD PEQUEÑA

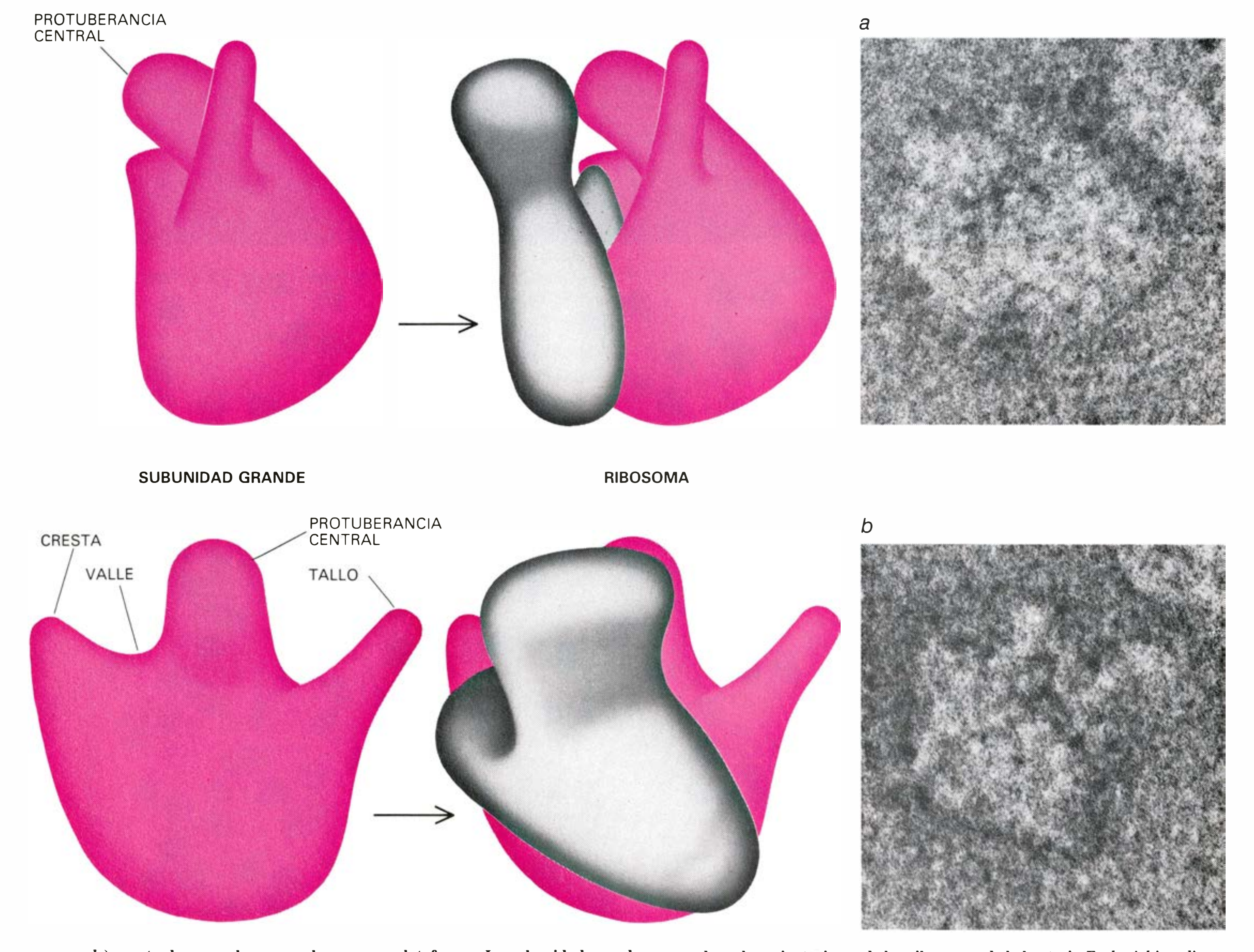


MODELO TRIDIMENSIONAL del ribosoma, con una configuración asimétrica en las dos subunidades que lo forman. La subunidad pequeña (*izquier-*

bía que "flotaban" libremente por el citoplasma; otros estaban unidos a membranas en el interior celular. Palade y Philip Siekevitz emplearon la ultracentrífuga para separar los diferentes constituyentes celulares de acuerdo con su tamaño. Una de las fracciones que obtuvieron (la fracción microsomal) mostraba abundancia de gránulos esféricos. Los análisis químicos revelaron que los gránulos eran ricos en ácido ribonucleico (ARN), hecho que condujo a la acuñación del término ribosoma para estos gránulos. Posteriormente se demostró que, si a la fracción se añadían aminoácidos marcados con átomos radiactivos, los aminoácidos (y su radiactividad) se incorporaban a las proteínas.

La calidad de las electromicrografías aumentó a lo largo de la década de los cincuenta, paralelamente a la mejora de las técnicas y los microscopios. En 1959, Henry S. Slayter y Cecil E. Hall, del Instituto de Tecnología de Massachusetts, y Geoffrey L. Zubay y H. E. Huxley, del Laboratorio de Biología Molecular del Consejo de Investigaciones Médicas de Cambridge, obtuvieron micrografías que probaban que el ribosoma estaba formado por dos subunidades. Medía una de ellas la mitad del tamaño de la otra, aproximadamente. Por entonces, la mayoría de los investigadores dudaban de que el microscopio electrónico pudiese dar mucha más información. Muchos pensaban que la estructura tridimensional del ribosoma sólo se determinaría cuando éstos pudiesen cristalizarse y ser sometidos a difracción de rayos X: proyecto virtualmente imposible para una estructura de la magnitud del ribosoma.

Con el camino hacia la difracción de rayos X aparentemente bloqueado, David DeRosier y Aaron Klug, del Laboratorio de Biología Molecular del Consejo de Investigaciones Médicas, volvieron a examinar las posibilidades que ofrecía el microscopio electrónico. En 1968, demostraron que una electromicrografía era matemáticamente equivalente a un patrón de difracción de rayos X. Específicamente, un patrón de difracción se relaciona con la estructura tridimensional de un cristal mediante la operación matemática conocida por transformación de Fourier. En cambio, la electromicrografía es un mapa que "reduce" una estructura tridimensional a dos dimensiones. La transformación de Fourier de los datos que componen el mapa, sin embargo, resulta (por el llamado teorema de proyección) en el plano central del patrón de difracción: el plano que pasa a través del centro del patrón y es perpendicular a la línea de visión a lo largo de la cual se ha hecho la electromicrografía. Tomando electromicrografías desde varias perspectivas se pueden componer los datos que habría facilitado la difracción de rayos X. La transformación de Fou-

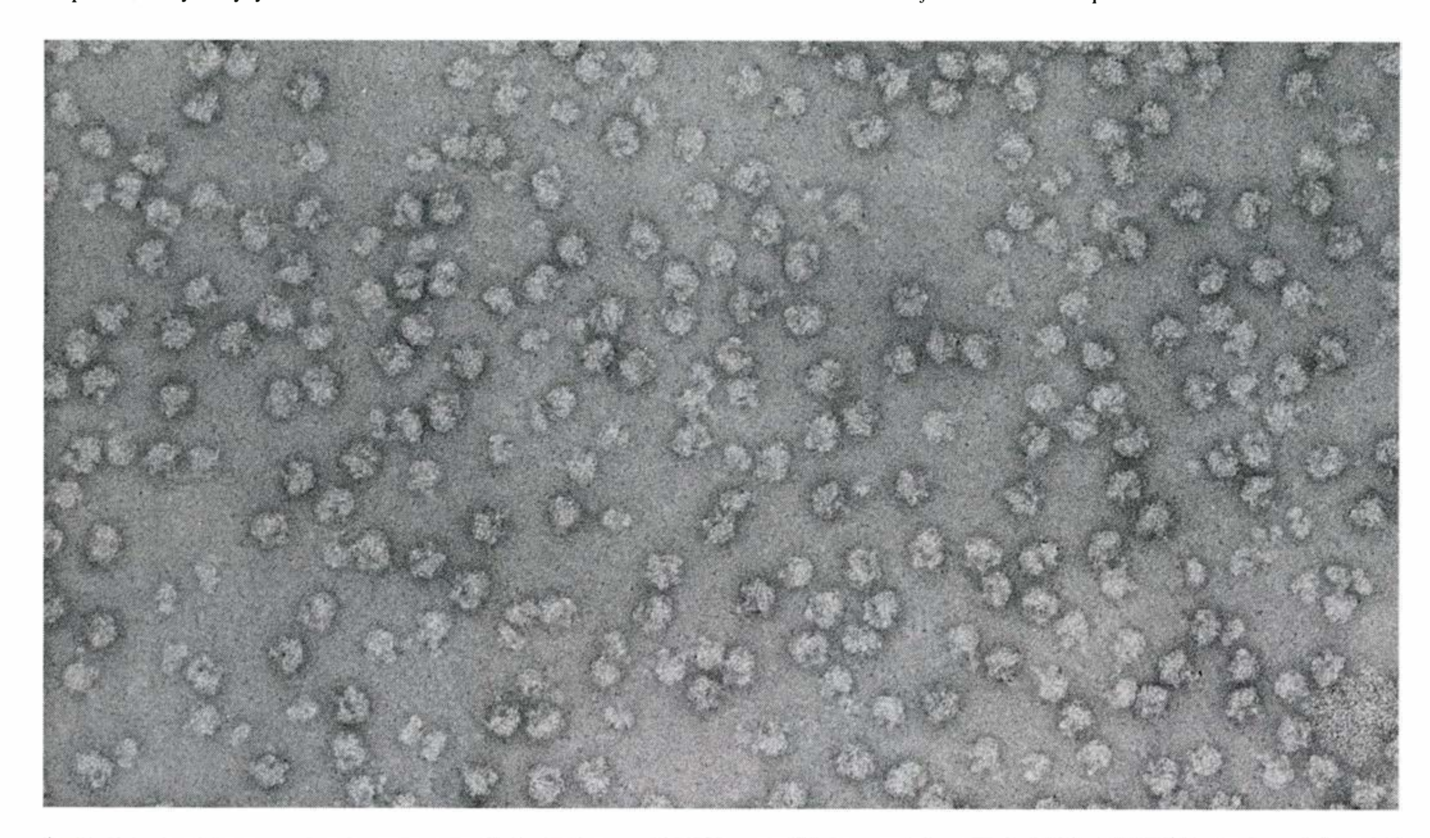


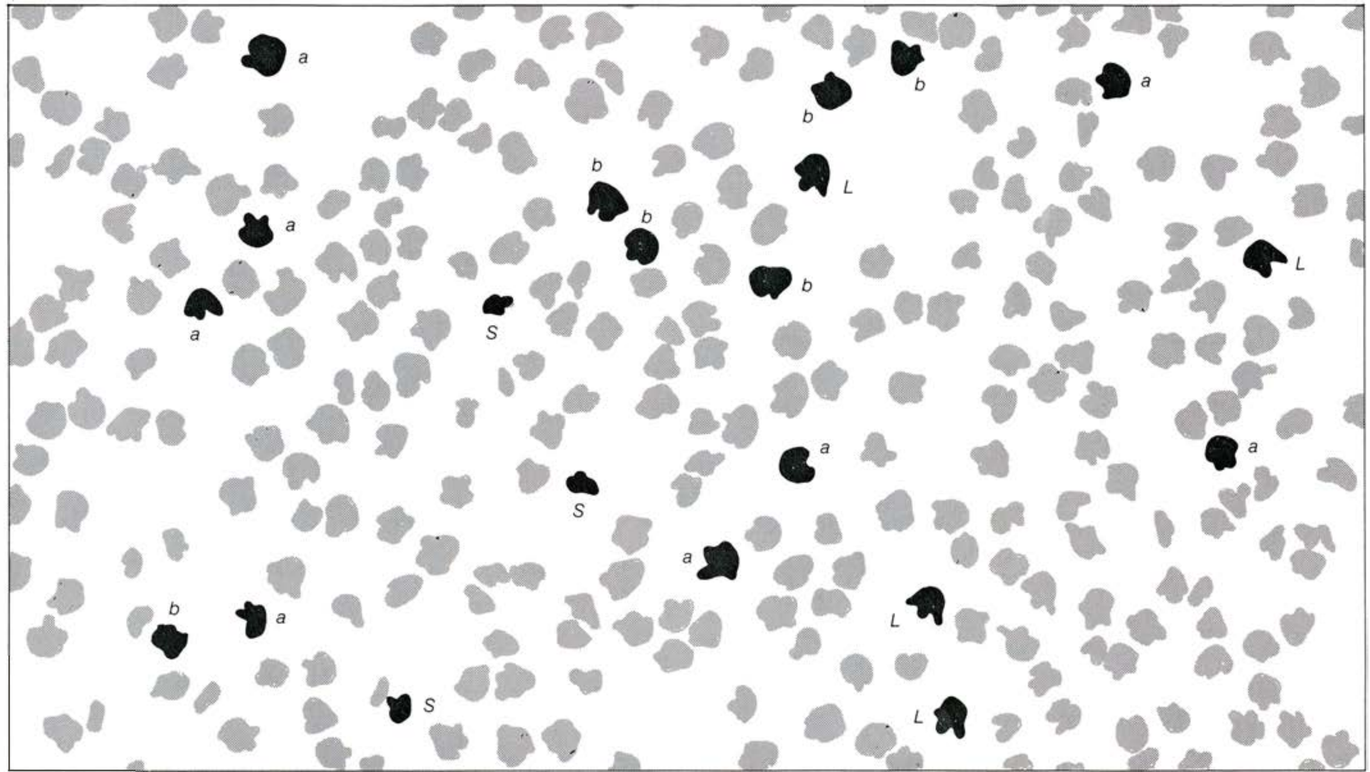
da) consta de una cabeza, una base y una plataforma. La subunidad grande (segunda por la izquierda) incluye una protuberancia central, flanqueada por una cresta y un tallo. Se muestran dos orientaciones del modelo, que corresponden a las orientaciones de los ribosomas de la bacteria *Escherichia coli* que aparecen en las electromicrografías de la derecha. El ribosoma tiene una longitud de 250 angstrom, aproximadamente. Las micrografías son del autor.

rier de estos datos origina posteriormente un mapa del patrón de densidad de toda la estructura tridimensional.

DeRosier y Klug utilizaron su técnica para producir, a partir de electromicrografías, un mapa de densidad tridimensional de parte de un virus. Con este impulso, Slayter y yo colaboramos, en la Facultad de Medicina de Harvard, en un esfuerzo por elaborar un mapa de densidad de un ribosoma. El resultado fue poco claro: la resolución del mapa era sólo de 100 angstrom. Aunque contentos por la imagen conseguida, nos sentimos decepcionados: el mapa no revelaba nada acerca de cómo trabaja el ribosoma. Era obvio que se necesitaba una detallada visión del ribosoma en tres dimensiones.

Los estudios bioquímicos del ribosoma también habían ido avanzando. La ultracentrífuga y otras técnicas separativas, fundamentalmente, habían demostrado que la menor de las dos subu-





PREPARACION DE RIBOSOMAS de *E. coli*; se observan algunas de las orientaciones presentadas en el modelo de ribosoma de las dos páginas precedentes. Las letras *a* y *b* los identifican en el mapa. Algunas subunidades ribosómicas, encontradas también en la preparación, se identifican con las letras *S* y *L*, que significan pequeña (*small*) y grande (*large*), respectivamente. La preparación para microscopía electrónica ha sido teñida negativamente; cada ribosoma está definido por un perfil de una sal de metal pesado, que en este caso es uranio. La electromicrografía fue obtenida, también, por el autor.

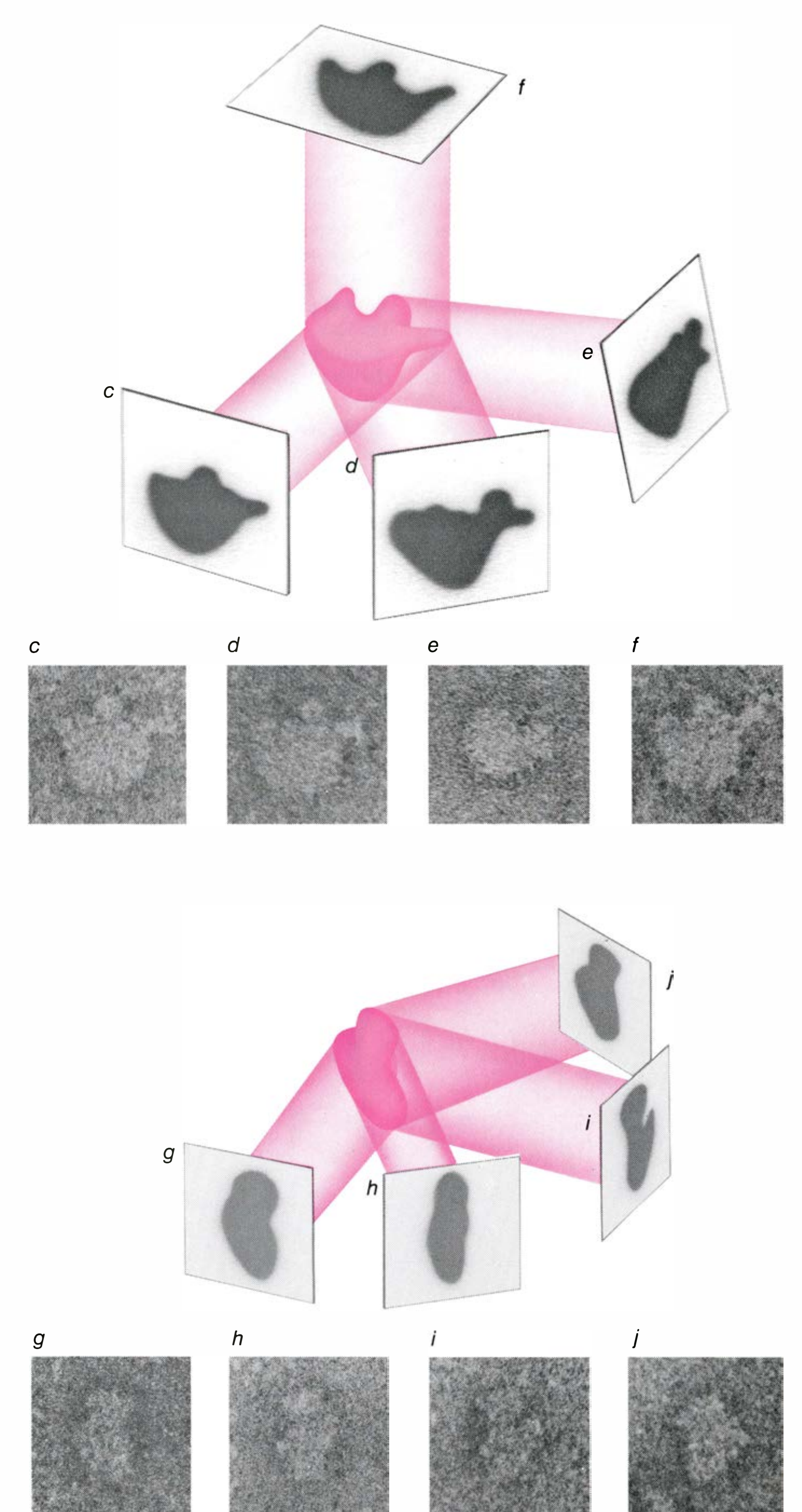
nidades del ribosoma estaba formada por una molécula de ARN, cuyo peso molecular era aproximadamente de 600.000, y 21 moléculas de proteína diferentes. El peso molecular total de la subunidad pequeña alcanzaba los 900.000. La subunidad mayor constaba de dos ARN y 34 proteínas. Su peso molecular total era aproximadamente de 1.600.000, y, de nuevo, las dos terceras partes del peso correspondían a ARN.

Técnicas experimentales

A finales de la década de 1960 se desarrolló una técnica mediante la cual el ARN y las proteínas constitutivas de la subunidad pequeña podían aislarse por separado y purificarse, de forma que, cuando estas diferentes sustancias se volvían a mezclar, se reconstruían las subunidades pequeñas. Las investigaciones iniciales habían sido hechas por Masayasu Nomura, Theophil Staehelin y Matthew Meselson, en la Universidad de Harvard. Finalmente, Nomura, en colaboración con Shoji Mizushima, de la Universidad de Wisconsin en Madison, realizó una serie de experimentos en los que las proteínas ribosómicas se añadían en diferente orden a una solución que contenía el ARN de la subunidad pequeña. Se proponía determinar qué proteínas debían unirse primero al ARN para que otras pudieran hacerlo después.

Se obtuvieron unos resultados complejos. Ninguna proteína inducía, por sí misma, en el autoensamblaje, la unión posterior de otra proteína a la subunidad pequeña. En cambio, las proteínas denominadas *S*4, *S*8, *S*15, *S*17 y *S*20 se unían directamente al ARN. La unión de *S*20 se facilitaba –esto es, ocurría con mayor eficacia– si *S*4 y *S*8 estaban ya unidas. La unión de cada proteína a la subunidad pequeña dependía en diferente grado de la presencia de más de una de sus antecesoras en el ensamblaje. La unión de la proteína *S*11, por ejemplo, se hacía más fácil si previamente se habían ya ensamblado no menos de ocho proteínas. Lo que resultaba del trabajo era un mapa con las secuencias de autoensamblaje de las proteínas.

El mapa de ensamblaje proporcionó varias pistas sobre las relaciones de vecindad entre las proteínas de la subunidad pequeña. Reflexiones posteriores sugirieron que tales pistas podían, en algunos casos, inducir a engaño. Cabe, por ejemplo, que la unión de una proteína al ARN de la subunidad pequeña pueda originar un cambio en la conformación del ARN, de suerte que un



EL MODELO DE RIBOSOMA SE DEDUJO a partir de diferentes imágenes de la subunidad grande (*arriba*) y pequeña (*abajo*), observadas en preparaciones de subunidades que se habían teñido negativamente para la microscopía electrónica. Las letras *c, d, e* y *f* identifican imágenes de la subunidad grande; *g, h, i* y *j*, imágenes de la subunidad pequeña. Cada imagen es la proyección bidimensional de una subunidad, que resulta de dirigir un rayo de electrones a través de una estructura tridimensional y la sal de uranio que la rodea. Las electromicrografías, que son de transmisión, no muestran la superficie del ribosoma; la configuración tridimensional del ribosoma se deduce a partir de múltiples proyecciones.

nuevo "bucle" de este ARN, quizás 100 angstrom más alejado del sitio de unión, se convierta ahora en el sitio de unión de una proteína que se incorpora más tarde. No obstante, la publicación del mapa de ensamblaje de Nomura renovó las esperanzas de desvelar algún día la estructura del ribosoma. Muy pronto, varios laboratorios pusieron a punto nuevas técnicas bioquímicas encaminadas a tal fin.

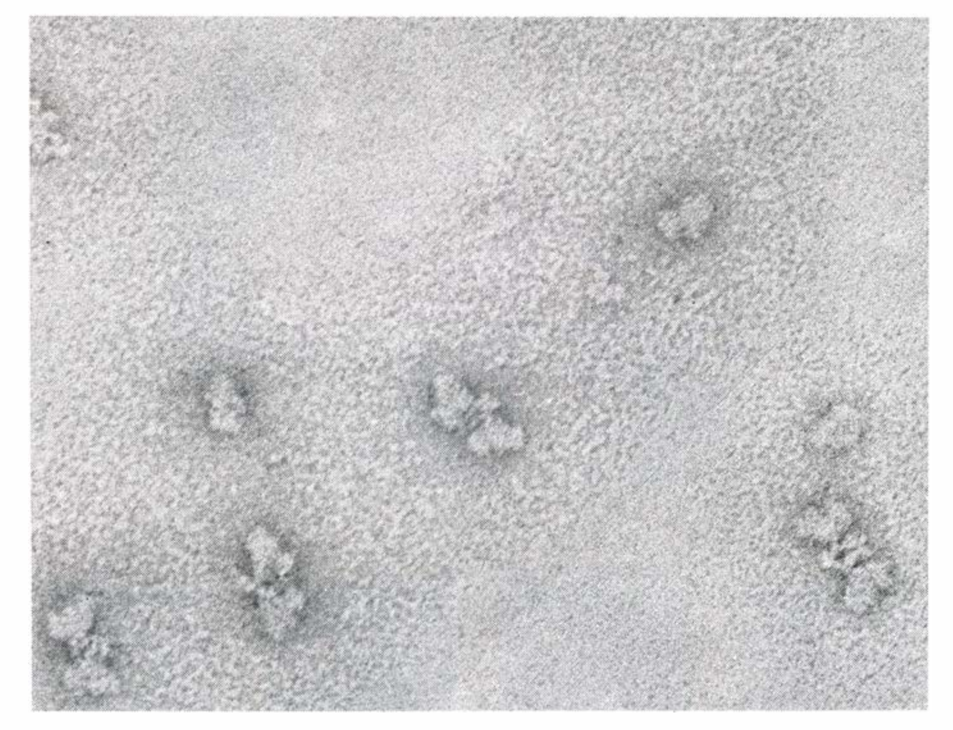
Y, entre ellas, la técnica de "soldadura", que mezcla los ribosomas con reactivos bifuncionales: moléculas que poseen dos sitios por los que se puede unir a otras moléculas. Una vez formados tales enlaces, se induce el "desensamblaje" de los ribosomas. Pueden entonces identificarse las parejas de proteínas que han quedado unidas por el reactivo. La técnica de soldadura se está empleando en muchos laboratorios para determinar qué proteínas ribosómicas son vecinas. Empeñados en ese tipo de investigación están los laboratorios de Robert R. Traut, en la Universidad de California en Davis, Charles G. Kurland, en la Universidad de Uppsala, Donal Hayes, en el Instituto de Química Física de París, y David Elson, en el Instituto Weizmann de Ciencias de Israel.

Una segunda técnica, ideada por Peter B. Moore y Donald M. Engelman, de la Universidad de Yale, requiere la preparación de proteínas de la subunidad pequeña en las que los átomos de hidrógeno se han reemplazado por átomos de deuterio (hidrógeno pesado). Se autoensamblan luego las subunidades pequeñas, a partir de una mezcla en la que dos de las proteínas presentes están deuteradas y el resto son normales. Las subunidades pequeñas se sitúan en un campo de neutrones de un reactor nuclear. El patrón de interferencia que resulta está dominado por la difracción de los neutrones por los átomos de deuterio. Lo que el patrón muestra es la distancia entre las dos proteínas deuteradas. A partir de varios experimentos como éste, en los que se preparan con deuterio diferentes parejas de proteínas de la subunidad pequeña, podemos obtener una serie de distancias. Finalizada la serie de distancias, se puede entonces determinar correctamente la posición del centro de masa de cada proteína.

En la década de 1970 se desarrolló con éxito una nueva técnica, la inmunomicroscopía electrónica. Está basada en la preparación de anticuerpos que se unen a proteínas ribosómicas específicas y permiten así revelar el lugar que estas proteínas ocupan en la superficie del ribosoma. El protocolo básico consiste en purificar una determinada proteína de un ribosoma bacteriano, por ejemplo, e inyectarla en la corriente sanguínea de un conejo u otro animal. La respuesta del sistema inmune del conejo a la presencia de la proteína ribosómica es la misma que se obtendría contra cualquier proteína de otra especie: los anticuerpos del conejo se unen a la sustancia extraña. Los anticuerpos pueden purificarse a partir de la sangre del conejo, y mezclarse con ribosomas enteros de la bacteria. Se espera que el anticuerpo se una a las partes de la proteína que están expuestas en la superficie del ribosoma.

En 1972, Georg Stöffler, del Instituto Max Planck de Genética Molecular de Berlín, Heinz-Günter Wittmann, del mismo centro, y Lawrence Kahan, que gozaba entonces de una beca postdoctoral en el laboratorio de Nomura, en Wisconsin, purificaron anticuerpos contra muchas proteínas de los ribosomas de la bacteria *Escherichia coli*. Cada vez estaba más claro que los anticuerpos podrían servir como marcadores para señalar proteínas específicas de la superficie del ribosoma. Nomura, Kahan y yo empezamos a trabajar en ensayos adicionales; Kahan siguió asociado con nosotros tras ingresar en el claustro docente de la Facultad en Wisconsin, y también William A. Strycharz, otro investigador postdoctoral del laboratorio de Nomura.

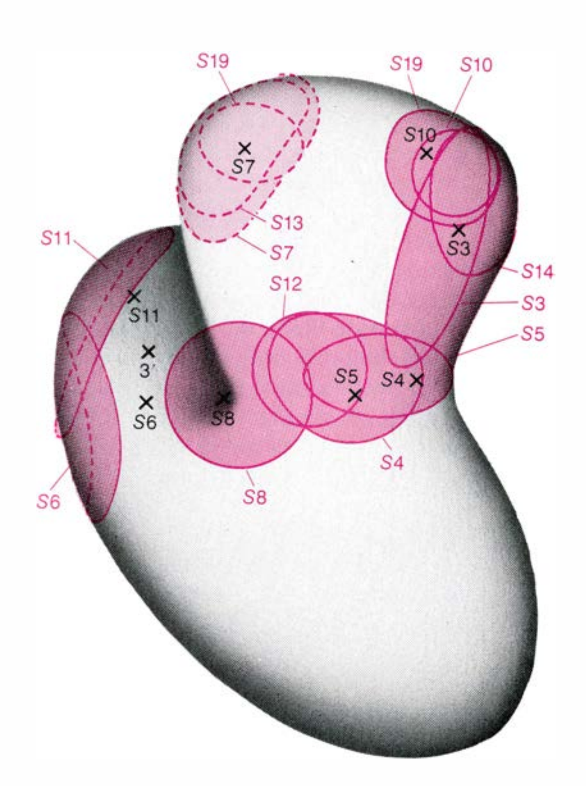
Nuestra estrategia de trabajo consistía en inducir en conejos la formación de los anticuerpos llamados inmunoglobulinas G (IgG) y hacerlos reaccionar con las subunidades pequeñas de los ribosomas. Los anticuerpos son moléculas en forma de *Y*, dotados de dos sitios de unión, en los extremos de cada uno de sus dos brazos cortos. Por tanto, el anticuerpo atrapa las subunidades a pares. Estas parejas pueden verse al microscopio electrónico. En general, su aspecto es variable, ya que poseen diferentes orientaciones. La localización del anticuerpo y sus lugares de unión a la subunidad pequeña permiten deducir no sólo la posición de la proteína, sino también la estructura tridimensional del ribosoma.

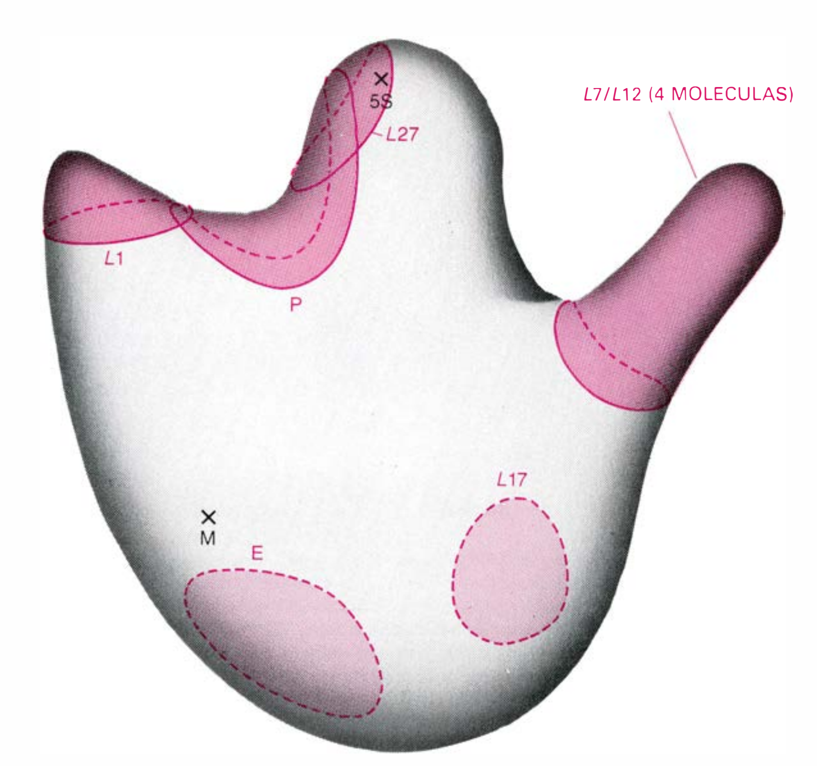


INMUNOELECTROMICROGRAFIA; revela el sitio (o sitios) que una proteína ribosómica ocupa en la superficie del ribosoma. Primero se purifica la proteína a partir de ribosomas bacterianos; después se inyecta en la corriente sanguínea de un animal de experimentación. El sistema inmune del animal fabrica anticuerpos contra la proteína. Los anticuerpos se mezclan con los ribosomas. En este caso se ha fabricado el anticuerpo inmunoglobulina G de conejo contra la proteína *S6* de la subunidad pequeña. Cada molécula de anticuerpo tiene forma de *Y*. Puede unir una *S6* en el extremo de cada uno de los brazos cortos de la *Y*. De esta forma, el anticuerpo "atrapa" las subunidades pequeñas a pares. El experimento fue realizado por el autor en colaboración con Lawrence Kahan, de la Universidad de Wisconsin.

Imágenes de transmisión

Llegados aquí, valdrá la pena examinar, con cierta detención, la naturaleza de la microscopía electrónica. La técnica que empleamos se llama de tinción negativa. Se cubre una rejilla horizontal de hilos de cobre con una capa de carbón. Sobre esta capa, se colocan los





MAPA DE LAS PROTEINAS ribosómicas; se muestra su localización en la subunidad pequeña (*izquierda*) y grande (*derecha*). Se han utilizado dos métodos. Las zonas coloreadas se cartografiaron por inmunomicroscopía electrónica; son los sitios de la superficie del ribosoma donde los anticuerpos se unen a las proteínas contra las que fueron fabricados. Cada proteína aparece una vez en la superficie de una subunidad, excepto la *S*19, que, como se ve, aparece en dos ocasiones; la proteína *S*4 está bajo la *S*5 y la *S*12. Las marcadas con una cruz se cartografiaron por difracción de neutrones, una técnica en la que las subunidades ribosómicas se "ensamblan" de forma que incluyan una pareja de proteínas en las que los átomos de hidrógeno se han sustituido por deuterio (hidrógeno pesado). El patrón de interferencia que se obtiene al dirigir un rayo de neutrones a través de la subunidad nos da la distancia que media entre las dos proteínas deuteradas. A partir de una serie de distancias de este tipo se deducen las posiciones de las proteínas. En la ilustración, las cruces señalan los centros de masa de distintas proteínas, determinados por difracción de neutrones. La inmunoelectromicroscopía fue realizada por el autor en colaboración con Lawrence Kahan, Masayasu Nomura y William A. Strycharz, en la Universidad de Wisconsin en Madison; la cartografía por difracción de neutrones se debe a Peter B. Moore y Donald M. Engelman, de la Universidad de Yale. Se muestran, además, cinco sitios adicionales. M es el sitio por el que el ribosoma se puede "enganchar" en la membrana intracelular (determinado por Nigel Unwin, de la Universidad de Stanford); 5*S*, situación de un ARN que forma parte de la subunidad grande (determinado por Vladimir Vasiliev y sus colaboradores en el Instituto de Proteínas de la Unión Soviética); 3', un extremo del ARN ribosómico denominado ARNr 16*S* (determinado por Helen Olson y Dohn G. Glitz, de la Universidad de California en Los Angeles); P, el sitio por donde se unen al ribosoma los sucesivos aminoácidos que formarán la cadena de proteína (determinado en el curso de los trabajos del autor), y E, el sitio por donde sale del ribosoma la cadena de proteína recién sintetizada (determinado por Carmelo Bernabev y el autor).

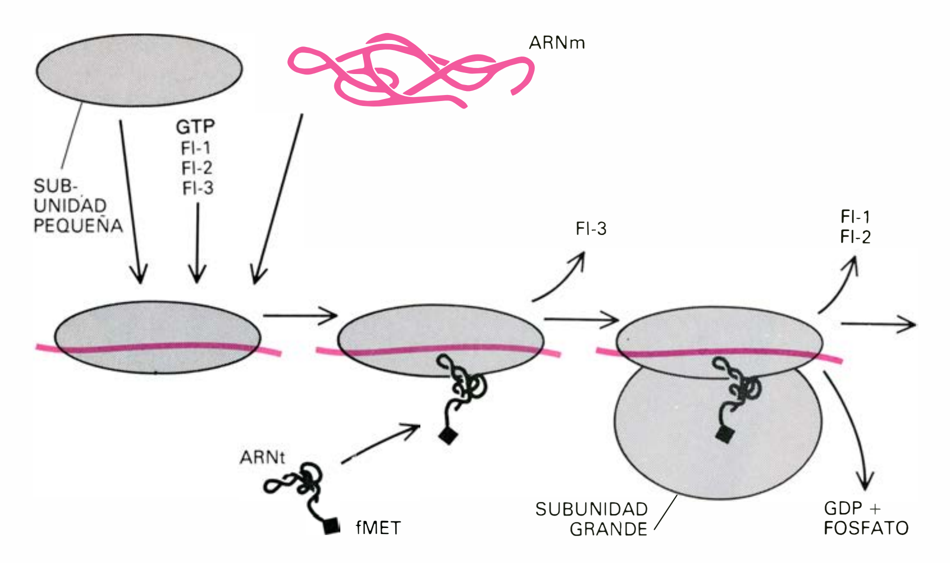
ribosomas (en solución). Se deposita luego una solución que contiene una sal de un metal pesado, como el uranio. Cuando los ribosomas se secan, cada uno de ellos se rodea de una película de átomos de uranio. Se introduce la preparación en el microscopio electrónico. Un rayo de electrones pasa sin mucha dificultad a través de los ribosomas. Es el uranio que rodea a cada ribosoma el que desvía los electrones y produce una zona más ensombrecida en la imagen resultante.

Una electromicrografía (como una placa de rayos X en medicina) es, fundamentalmente, una imagen transmitida. Se trata de la proyección bidimensional de una estructura tridimensional que resulta de dirigir un rayo a través de la estructura. Una fotografía normal es, por contra, una grabación de luz reflejada. En otras palabras, una fotografía normal muestra superficies; no así una imagen de transmisión. Cuando se estudia una imagen de transmisión, la experiencia visual propia es menos útil de lo que uno desearía.

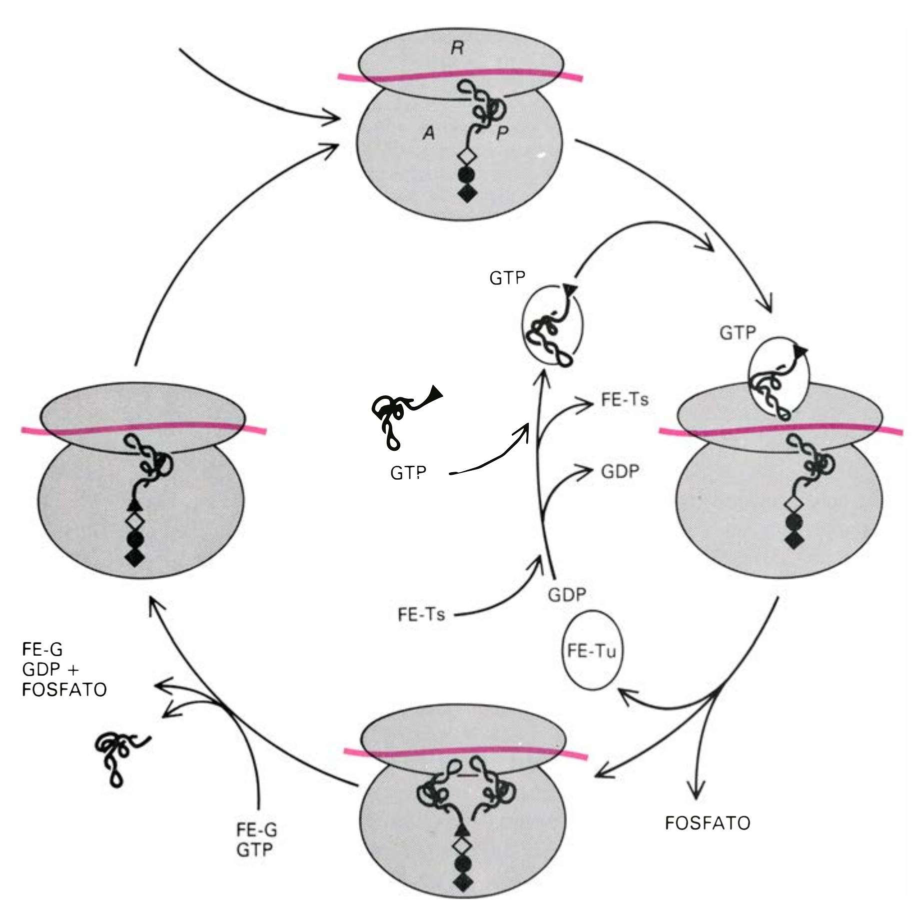
Por ejemplo, en una imagen transmitida no hay forma de distinguir qué parte de una estructura está más cerca y cuál más lejos. Imagine que le enseñan una placa de rayos X de una mano. Los dedos se extienden hacia la parte superior de la imagen y el pulgar está a la derecha. ¿Corresponde esa radiografía a la mano derecha o a la izquierda? No puede responderse a esta pregunta. La placa de rayos X de una mano izquierda con la palma hacia abajo es indistinguible de la radiografía de una mano derecha con la palma hacia arriba. Esta propiedad de las imágenes de transmisión tiene una implicación inmediata en las electromicrografías de una preparación de ribosomas orientados al azar en una rejilla de hilos de cobre. Si aparece una determinada imagen de un ribosoma, un enantiomorfo, o imagen especular, probablemente esté en alguna parte de la preparación.

Las proyecciones del ribosoma no son más que indicios acerca de su conformación tridimensional. Pero podemos esperar hallar varias proyecciones distintas del ribosoma en las preparaciones teñidas negativamente y deducir la conformación que produce estas proyecciones. Y eso fue lo que, en 1970, Yoshiaki Nonomuta, Günter Klaus-Joachim Blobel y David D. Sabatini realizaron en la Universidad Rockefeller. Identificaron perfiles específicos en las electromicrografías de ribosomas de células hepáticas de rata y dedujeron un modelo de baja resolución de una posible estructura tridimensional.

Surge un segundo problema cuando se trata de obtener un modelo de alta resolución. La proyección característica de cada subunidad ribosómica es más o menos especularmente simétrica. Por tanto, un modelo tridimensional deducido a partir de dos proyecciones (la proyección cuasisimétrica y una segunda perspectiva distintiva) tendrá simetría bilateral. El hecho de que las electromicrografías sean imágenes de transmisión determina que cualquier inmunoelectromicrografía que se interprete para situar una proteína en el la-



INICIACION, primera etapa en la síntesis de una proteína; requiere que varias moléculas se unan a la subunidad pequeña del ribosoma. Concretamente, se unen tres factores de iniciación (FI) junto con trifosfato de guanosina (GTP) y ARN mensajero (ARNm). El GTP es una fuente de energía; perderá uno de sus grupos fosfatos para convertirse en difosfato de guanosina (GDP). El ARNm es una cinta de instrucciones; su secuencia de unidades, las bases, constituye un código para la secuencia de aminoácidos que forman una proteína. A continuación, (en bacterias) el aminoácido modificado formilmetionina (fMet) se incorpora a la subunidad pequeña traído por el ARN de transferencia (ARNt), una molécula especializada para unir en uno de sus extremos un aminoácido específico y, por el otro, unirse al ribosoma. El complejo que resulta de la iniciación consta de una subunidad ribosómica grande, una pequeña, ARNm y ARNt-fMet.



CICLO DE ELONGACION, segunda etapa en la síntesis de una proteína; añade aminoácidos a la cadena naciente (parcialmente sintetizada) de proteína. Cada repetición del ciclo requiere la participación de las moléculas denominadas factores de elongación (FE) y también de dos moléculas de GTP. Al comienzo del ciclo, un ARNt peptidílico (un ARNt que lleva la cadena naciente) se une al ribosoma en el sitio denominado *P*. Un ARNt aminoacílico (un ARNt que lleva el aminoácido determinado por el codón que se leerá a continuación en el ARNm) se une al sitio *R*. Y lo hace junto con FE-Tu y GTP. La unión está controlada por la de un codón del ARNm a tres bases ("anticodón") del ARNt. El ARNt se transfiere al sitio *A*, donde su aminoácido incorpora la cadena naciente. Finalmente, se expulsa el ARNt del sitio *P* y el ARNt del sitio *A* ocupa su lugar. Otro ciclo prepara el complejo formado por FE-Tu, GTP y un ARNt aminoacílico.

do izquierdo de un modelo simétrico puede interpretarse con igual justificación para situarla en el lado derecho. Me sentía optimista por haber encontrado tres proyecciones distintas de la subunidad pequeña. Una de ellas era la proyección cuasisimétrica. En ese caso, sin embargo, la estructura tridimensional sería completamente asimétrica. Los sitios de unión de varias proteínas de la subunidad pequeña, en cada una de las proyecciones (y en sus enantiomorfos), constituían puntos de referencia para relizar las rotaciones que relacionaban las tres proyecciones. Tardé varios meses en determinar un modelo tridimensional asimétrico.

Otros dos grupos habían obtenido también un modelo. Examinando electromicrografías de subunidades pequeñas teñidas negativamente, Stöffler, auxiliado por Gilbert Tischendorf, un estudiante graduado de su laboratorio, consiguió un modelo simétrico. Vladimir Vasiliev, del Instituto de Proteínas de la Unión Soviética, siguió una estrategia diferente. Cubrió las preparaciones de subunidades pequeñas con metal pesado, "bombardeándolas" con los átomos de metal desde un ángulo, de forma que el metal se depositara alrededor de cada ribosoma como si fuese una tormenta de nieve. Aunque las electromicrografías podían tratarse entonces como imágenes reflejadas, tal "sombreado" producía una resolución inferior a la de la tinción negativa. En efecto, la resolución era a menudo demasiado baja para revelar la localización de los anticuerpos empleados en la inmunomicroscopía electrónica. Vasiliev, no obstante, obtuvo un modelo asimétrico de la subunidad pequeña del ribosoma que se parecía mucho al nuestro.

Una vez alcanzado un modelo asimétrico, para definir sin ambigüedad la posición del sitio de unión de una proteína en la superficie tridimensional de la subunidad pequeña bastan inmunoelectromicrografías que muestren dos perspectivas distintas de dicho sitio de unión. Cada perspectiva determina una línea a través del ribosoma, sobre la que debe hallarse el sitio de unión. La intersección de las líneas establece el sitio. Nuestra cartografía de proteínas continuaba. Mientras tanto, en 1979, Moore, Engelman y sus colaboradores habían cartografiado suficientes proteínas de la subunidad pequeña por difracción de neutrones. Había llegado el momento de comparar sus resultados con los nuestros. Sus datos se ajustaban al modelo asimétrico y no al simétrico.

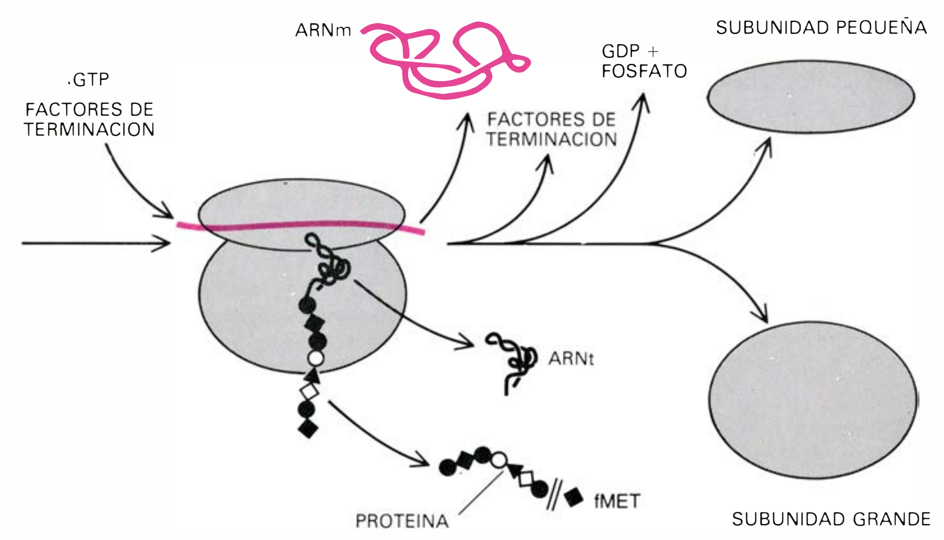
Para facilitar el análisis de la estruc-

tura asimétrica del ribosoma, le hemos asignado nombres característicos. El modelo de la subunidad pequeña consta de tres regiones: una cabeza, que constituye aproximadamente un tercio de la subunidad pequeña; una base, que representa los otros dos tercios, y una plataforma, que está separada de la cabeza por una cavidad que denominamos hendidura. Una constricción en la subunidad pequeña delimita la cabeza de la base. El modelo de la subunidad grande (obtenido por métodos similares) incluye una protuberancia central flanqueada por una proyección a cada lado. La proyección más alargada es el tallo y, cresta, la otra proyección. Entre la cresta y la protuberancia central se sitúa el valle. Cuando las subunidades grande y pequeña se unen, el tallo de la subunidad grande tiene su base cerca de la constricción de la subunidad pequeña, y la cabeza de la subunidad pequeña y la protuberancia central de la subunidad grande están aproximadamente alineadas.

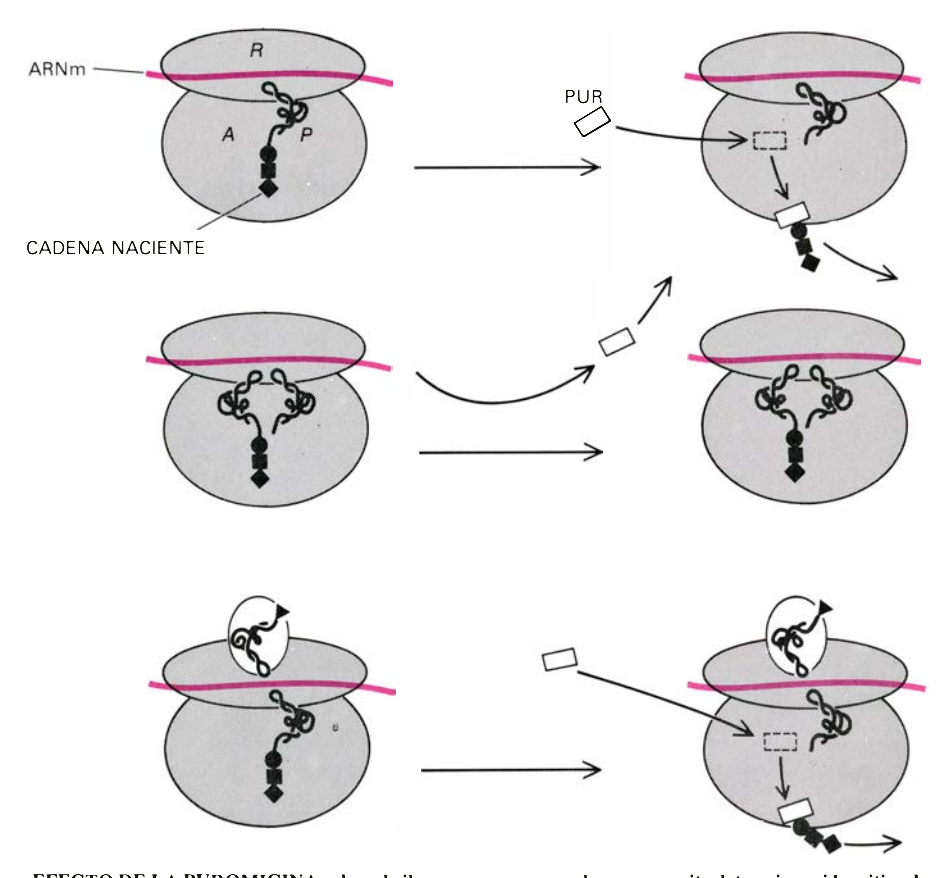
Proteínas de fidelidad

Los estudios bioquímicos sobre esta estructura están empezando a dar sus frutos. Consideremos la proteína *S*4. Durante el análisis cartográfico de las proteínas de la subunidad pequeña, mediante inmunomicroscopía electrónica, Kahan, trabajando con Donald A. Winkelmann, un estudiante graduado en Wisconsin, descubrió que el anticuerpo que había preparado sensibilizando el sistema inmune del conejo a la proteína *S*4 de *E. coli* podía unirse a la propia proteína *S*4 y al complejo formado por la unión de la proteína *S*4 con el ARN de la subunidad pequeña. Sin embargo, no podía unirse a las subunidades pequeñas. Kahan y Winkelmann probaron la reactividad del anticuerpo mezclándolo con subunidades pequeñas en diferentes grados de ensamblaje. Encontraron que el anticuerpo contra la proteína *S*4 sólo se unía a la subunidad pequeña si faltaban las proteínas *S*5 y/o *S*12. Es como si *S*5 y *S*12 "tapasen" a *S*4. En efecto, cuando recientemente cartografiamos los sitios de unión de la proteína *S*4 en subunidades sin *S*5 ni *S*12, encontramos que el sitio *S*4 estaba situado donde se hallaría el de *S*5 si la subunidad estuviese completa.

La proteína *S*4 es poco corriente en muchos sentidos. Se une fuertemente al ARN de la subunidad pequeña, se une al ARN en ausencia de cualquier otra proteína ribosómica y, una vez enlazada, protege a una gran parte del ARN de la degradación por los enzimas llamados nucleasas. Además, la proteína *S*4 ocupa una posición clave en el mapa de ensamblaje de Nomura, colaborando en la unión de otras cuatro proteínas. Pero se trata de una colaboración débil. Además, nuestra cartografía de proteínas no sugiere una relación simple entre las interacciones de ensamblaje de *S*4 y su localización. Por ejemplo, *S*4 no tiene interacciones de ensamblaje directas con *S*5 y *S*12, aun cuando las tres estén situadas juntas en la cara de la subunidad pequeña que se halla más alejada de la subunidad grande.



TERMINACION. Acaba la síntesis de una proteína y ésta sale del ribosoma. También sale el ARNm. La siguiente vez que el ribosoma empiece a sintetizar una proteína, deberá ensamblar de nuevo las subunidades. La terminación requiere dos factores de terminación y, probablemente, una molécula de GTP.



EFECTO DE LA PUROMICINA sobre el ribosoma; es una prueba que permite determinar si los sitios de unión de los ARNt están ocupados. Al comenzar una repetición del ciclo de elongación (*línea superior*), un ARNt peptidílico ocupa el sitio *P*. El antibiótico puromicina (Pur) puede ocupar parte del sitio *A*, incorporar la cadena naciente de proteína y a continuación "desprenderla" del ribosoma. A mitad del ciclo (*línea central*) el sitio *A* está ocupado por un ARNt peptidílico, y la molécula de puromicina que llega no se puede unir. En condiciones experimentales se puede inmovilizar un ARNt aminoacílico en el sitio *R* mientras el *P* está ocupado por un ARN de transferencia peptidílico (*línea inferior*). La puromicina, entonces, hace que se libere la cadena naciente. Se deduce que el sitio *A* está vacío cuando el sitio *R* está ocupado.

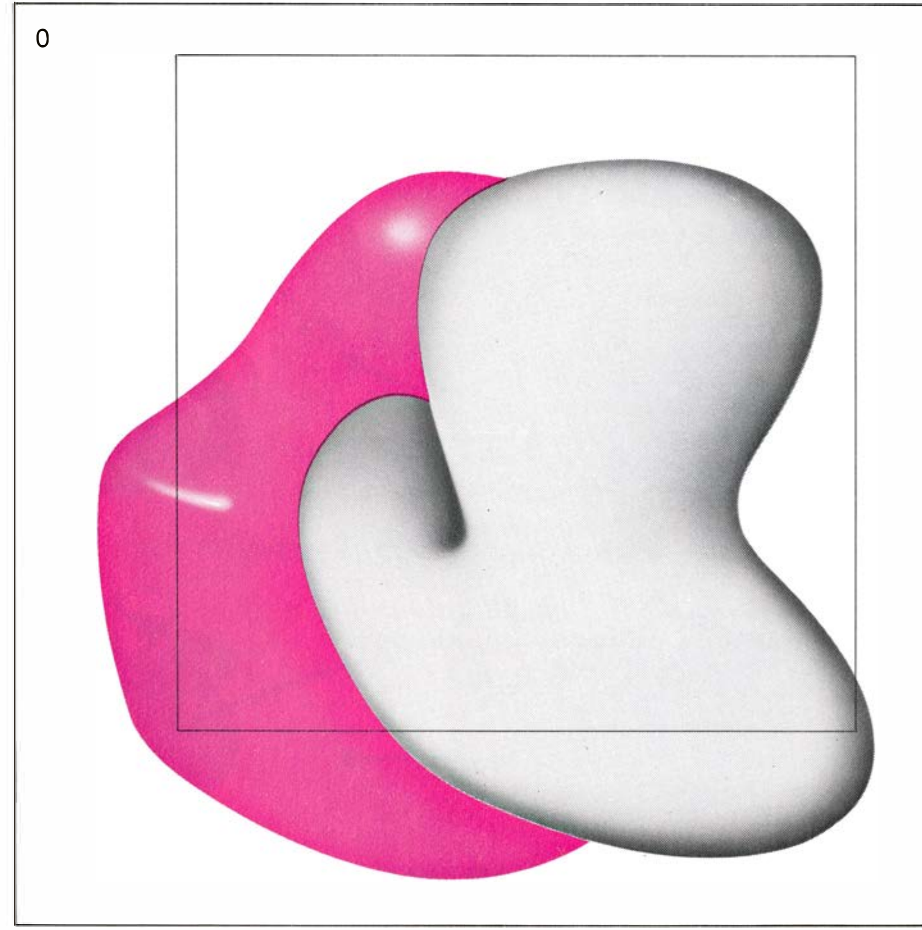
*S*4, *S*5 y *S*12 también están ligadas funcionalmente. En particular, se sabe que el antibiótico estreptomicina hace que el ribosoma fabrique proteínas llenas de errores en su secuencia de aminoácidos. Los ribosomas de ciertas estirpes mutantes de bacterias resisten este efecto de la droga. Si se purifica la proteína *S*12 de los ribosomas de bacterias resistentes y se introduce en vez de la *S*12 presente en los ribosomas de bacterias normales, éstas se tornan resistentes. Experimentos con otras drogas que determinan "mala lectura" (y por tanto errores) también implican a *S*4 y *S*5 como "proteínas de fidelidad".

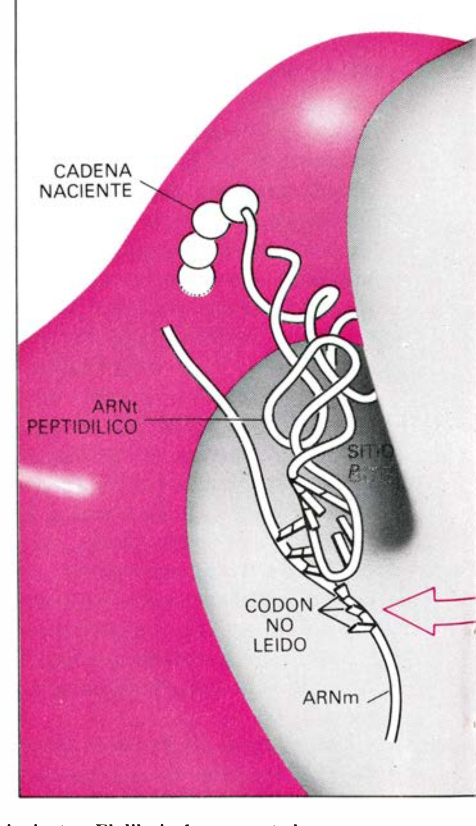
Bioquímica de la síntesis de proteínas

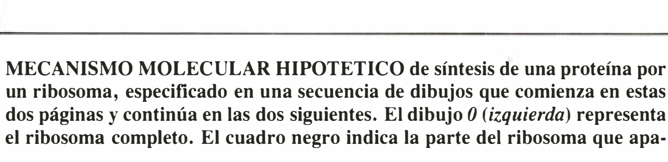
La localización de las proteínas *S*4, *S*5 y *S*12 en un lugar específico de la subunidad pequeña, junto con los descubrimientos bioquímicos sobre la función de las proteínas, constituyen un ejemplo de cómo determinadas regiones del ribosoma están empezando a adquirir relevancia. Antes de entrar en estos aspectos, sin embargo, hemos de decir algo más sobre la síntesis de proteínas, entendida como proceso bioquímico y como consumidor de los recursos energéticos de la célula.

La síntesis de una proteína se desarrolla a lo largo de tres etapas diferentes. La primera etapa es la de iniciación: en ella, el ARN mensajero (ARNm), que se sintetizó sobre un molde de ADN, se sitúa de tal forma en el ribosoma que su mensaje se puede traducir secuencialmente en una cadena de unidades de aminoácidos. El proceso comienza cuando tres moléculas de proteínas, llamadas factores de iniciación (FI), se unen a la subunidad pequeña del ribosoma. También se unen otras dos moléculas: trifosfato de guanosina (GTP) y ARNm. El trifosfato de guanosina es una fuente de energía. Tiene tres enlaces ricos en energía que unen los grupos fosfato (PO_3) a la guanosina, y, al final de un proceso en el que el GTP contribuye energéticamente, éste se convierte en difosfato de guanosina (GDP) y un fosfato inorgánico. El ARNm contiene las instrucciones genéticas. Las instrucciones (el mensaje) consisten en una secuencia de bases que sobresalen de la cadena de ARN como una pila de baldosas. Un triplete de bases (codón) determina un aminoácido particular.

En las bacterias, el codón "iniciador" (triplete de bases situado al principio del mensaje transcrito) es adenina, uracilo, guanina (AUG) o, a veces, guanina, uracilo, guanina (GUG). Como codón iniciador, cualquiera de las dos secuencias determina el aminoácido metionina. La molécula situada por el ribosoma al comienzo de una cadena proteínica bacteriana es una versión modificada de la metionina, conocida como formilmetionina (fMet). La iniciación, que incluye la traducción del codón iniciador, requiere por tanto que la fMet esté unida a la subunidad pequeña. El aminoácido es llevado hasta allí por un ARN de transferencia (ARNt), una molécula especializada que, por uno de sus extremos, se puede unir a un aminoácido particular y, por el otro, al ribosoma. La unión del ARNt al ribosoma provoca la salida de un factor de iniciación, el FI-3. El agregado que queda se une entonces a la subunidad grande. Inmediata-






MECANISMO MOLECULAR HIPOTETICO de síntesis de una proteína por un ribosoma, especificado en una secuencia de dibujos que comienza en estas dos páginas y continúa en las dos siguientes. El dibujo *0* (*izquierda*) representa el ribosoma completo. El cuadro negro indica la parte del ribosoma que aparece en cada uno de los dibujos siguientes. El dibujo *1* representa los sucesos que tienen lugar al comienzo de una repetición del ciclo de elongación. Un ARNt peptidílico está unido al sitio *P* de la subunidad pequeña; lleva una cadena naciente de proteína que se representa en el momento de entrar en la

mente, el GTP, disociado en GDP y fosfato inorgánico, se desprende. Después lo hacen FI-1 y FI-2.

La segunda etapa de la síntesis de una proteína se denomina elongación. Se trata de un ciclo que se repite cada vez que un aminoácido, desde el segundo hasta el último, se añade a la cadena naciente (parcialmente sintetizada) de proteína. (El segundo aminoácido está, por supuesto, unido a fMet.) Cada repetición del ciclo de elongación requiere que dos proteínas, llamadas factores de elongación, se unan temporalmente al ribosoma. También requiere que dos moléculas de GTP se unan al ribosoma y se disocien en GDP y fosfato inorgánico.

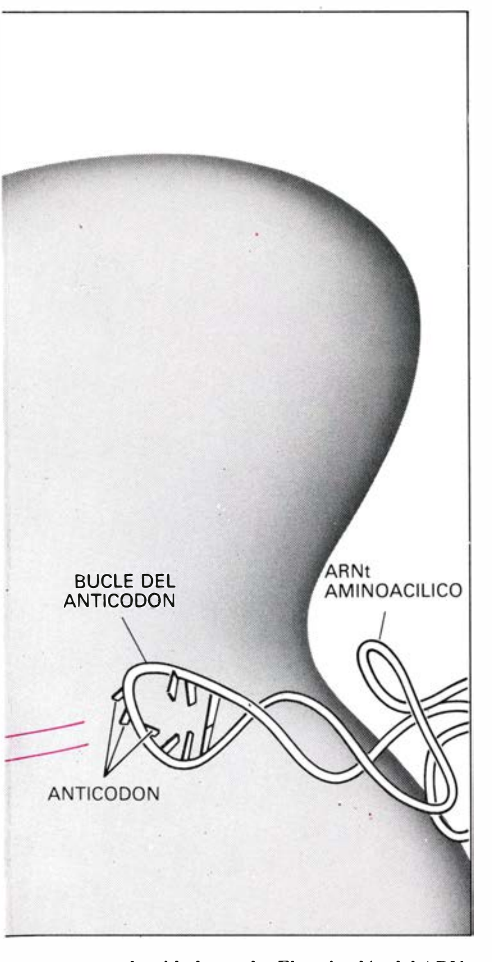
La etapa final de la síntesis de proteínas no se conoce tan bien como las dos anteriores. Están involucradas dos proteínas llamadas factores de terminación, y parece que UAG, UAA y UGA sirven todos como codones terminadores: tripletes de ARNm que determinan que el ribosoma libere al mensajero y a la proteína recién sintetizada. La terminación requiere, muy probablemente, que una molécula de GTP se una al ribosoma y se disocie en GDP y fosfato.

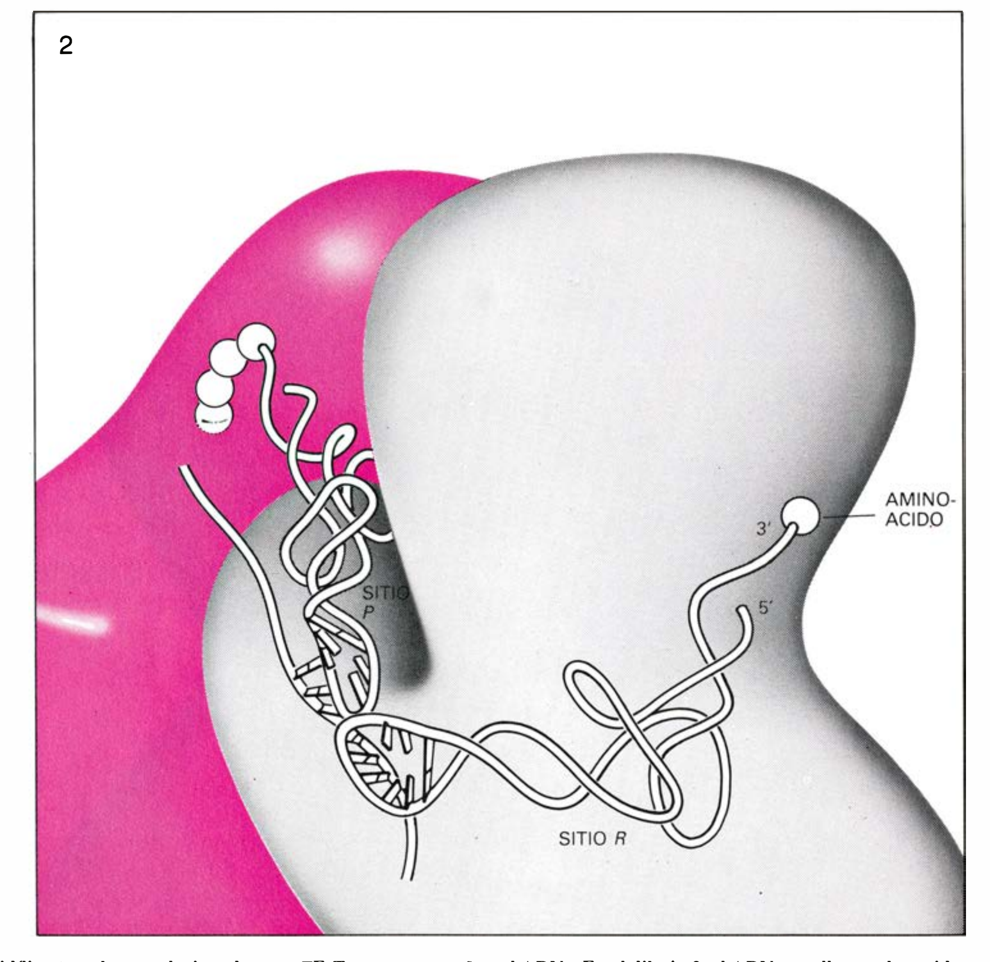
Ciclo de elongación

Los esfuerzos realizados para desvelar los pormenores de la síntesis de proteínas se han centrado principalmente en la segunda etapa: el ciclo de elongación. En este sentido, el antibiótico puromicina ha desempeñado un papel crucial. La puromicina es una molécula capaz de interrumpir la síntesis de proteínas en ciertos momentos del ciclo de elongación. Su estructura mimetiza el extremo del ARN transferente que lleva un aminoácido. El mimetismo es tan fiel que, cuando la puromicina ha usurpado el sitio de unión del ARNt, el ribosoma erróneamente la une a la cadena naciente de proteína. La unión de la puromicina al ribosoma resulta ser más débil de lo que hubiera sido la de un ARNt. La cadena incompleta, terminada ahora con puromicina, "cae" del ribosoma, y el ciclo de elongación se para.

Durante 15 años se ha venido sacándoles partido a las propiedades de la puromicina para establecer que los ARNt tienen dos sitios de unión en el ribosoma. En efecto, el modelo de ciclo de elongación contempla que los ARNt se están transportando sucesivamente entre los dos sitios. Al comienzo de cada repetición del ciclo, un ARNt peptidílico (un ARNt que lleva la cadena naciente de proteína) se une al llamado sitio *P*. Un ARNt aminoacílico (un ARNt con un aminoácido) se une al sitio *A*. Este aminoácido es el determinado por el codón que debe leerse a continuación. La cadena naciente se separa del ARNt en el sitio *P* y se une al aminoácido del ARNt recién llegado. El ARNt del sitio *P*, despojado ahora de la cadena naciente, es expulsado del ribosoma; el ARNt del sitio *A* ocupa ahora su lugar, dejando el sitio *A* vacío para la siguiente repetición del ciclo.

Los efectos de la puromicina están de acuerdo con este esquema. Al comienzo del ciclo, cuando un ARNt que lleva la cadena naciente ocupa el sitio *P*, la puromicina puede ocupar el sitio *A* y determinar posteriormente la "caída" de la cadena naciente del ribosoma. A mitad del ciclo, cuando un ARNt con la





subunidad grande. El anticodón del ARNt peptidílico (tres bases relacionadas con el aminoácido al que va unido el ARNt) está unido al codón del ARNm, que se tradujo en el ciclo anterior. Cerca de la subunidad pequeña se sitúa un ARNt aminoacílico. No se representan ni el GTP ni el factor de elongación FE-Tu que acompañan al ARNt. En el dibujo *2*, el ARNt que llega se ha unido al sitio *R*; su anticodón se une al codón que se va a leer a continuación. El anticodón forma parte de un bucle de siete bases. Cinco de ellas, ahora las situadas más cerca del extremo del ARNt denominado 5', están "apiladas".

cadena naciente ocupa el sitio *A*, la puromicina no puede unirse al ribosoma. No puede interrumpir, por tanto, la síntesis de la proteína. La introducción de la puromicina constituye una prueba para saber si el sitio *A* está o no ocupado.

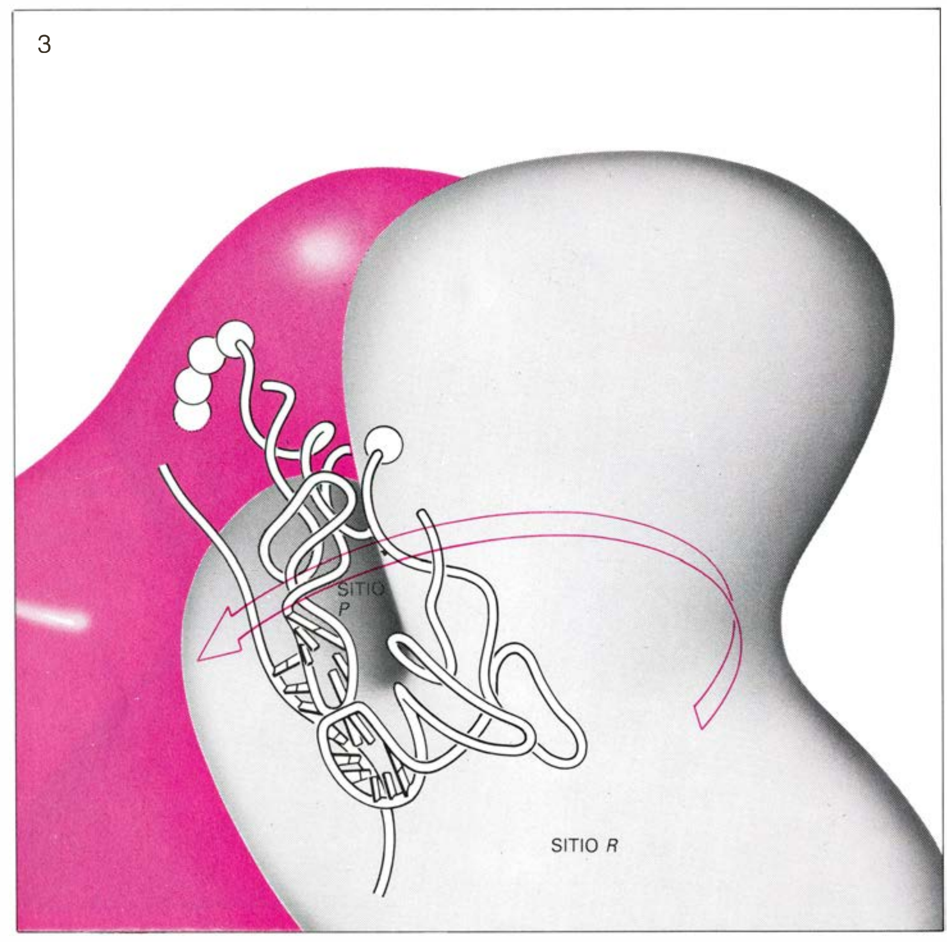
A principios de la década de 1970 se hizo evidente que este modelo era incompleto. ¿Por qué cada repetición del ciclo de elongación requería la disociación de dos moléculas de GTP? La disociación de la segunda molécula es fácil de explicar. Puede suministrar la energía necesaria para mover al ARNt peptidílico desde el sitio *A* hasta el *P*, y hacer avanzar el ARNm para que pueda leerse el siguiente codón. Cuesta más justificar la disociación de la primera molécula; en particular, porque la energía liberada al romperse el enlace que une al ARNt peptidílico con la cadena naciente es más que suficiente para unir la cadena naciente al aminoácido del ARNt aminoacílico.

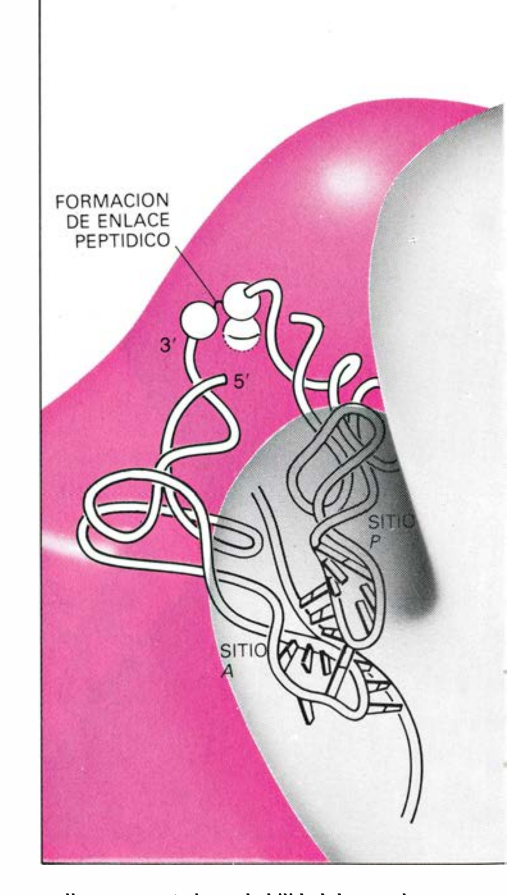
En efecto, por un experimento realizado en 1969 por Jean M. Lucas-Lenard, de la Universidad Rockefeller, podemos ver ahora que el modelo resultaba demasiado "estrecho". Unos años más tarde el resultado de este experimento fue confirmado por Yoshito Kaziro y sus colegas, de la Universidad de Tokio. El experimento está basado en el hecho de que, cuando un ARNt aminoacílico se une al ribosoma, lo hace en combinación con otras dos moléculas: el factor de elongación llamado FE-Tu y también una molécula de GTP. En el tubo de ensayo se puede sustituir el GTP por una sustancia llamada GMPPCP. Es idéntica al GTP, salvo en que un átomo de carbono ocupa el lugar de un átomo de oxígeno. El ARNt puede unirse al ribosoma en combinación con FE-Tu y GMPPCP. La molécula de GMPPCP, sin embargo, no puede disociarse. El enlace fosfato, por tanto, no suministra energía y el ARNt no puede moverse del lugar donde está unido.

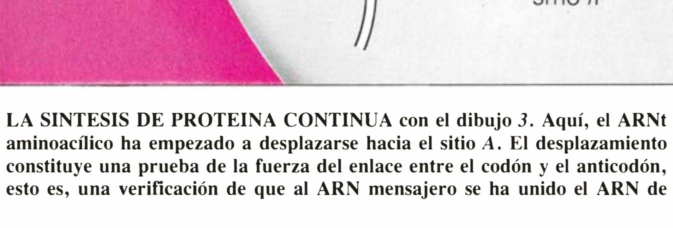
Lucas-Lenard fabricó el complejo consistente en un ARNt aminoacílico, FE-Tu y GMPPCP, lo dejó unirse al ribosoma a la mitad del ciclo de elongación y demostró que la puromicina podía liberar la cadena naciente. En ese momento, el sitio *P* estaba ocupado por el ARNt peptidílico y el sitio *A*, vacío. (Esto se dedujo porque el sitio *A* aceptaba la puromicina.) Con todo, el ARNt aminoacílico había encontrado un sitio para unirse. En la primavera de 1977 propuse en una publicación la existencia de un tercer sitio de unión para el ARNt. Lo llamé sitio *R*. Seis meses más tarde, Arthur Johnson, a la sazón en la Facultad de Medicina y Cirugía de la Universidad de Columbia, publicó, independientemente, una propuesta similar.

ARN de transferencia

Para entender la función que puede cumplir el sitio *R*, es preciso conocer la estructura básica de un ARN de transferencia, tal como se obtiene por difracción de rayos X. Al igual que ocurre con otras moléculas de importancia biológica, podemos caracterizar el ARN de transferencia describiendo su estructura primaria, secundaria y terciaria. La estructura primaria es simplemente la composición lineal de la molécula: una hilera de unas 80 bases nucleotídicas, que sobresalen de un esqueleto de azúcar y fosfato. Tres de estas bases (guanina, citosina y adenina) también se encuentran en el ADN. Pe-






LA SINTESIS DE PROTEINA CONTINUA con el dibujo *3*. Aquí, el ARNt aminoacílico ha empezado a desplazarse hacia el sitio *A*. El desplazamiento constituye una prueba de la fuerza del enlace entre el codón y el anticodón, esto es, una verificación de que al ARN mensajero se ha unido el ARN de transferencia que corresponde; con ello se aumenta la probabilidad de que el aminoácido que se incorpora a la proteína sea el correcto. En el dibujo *4*, el ARNt aminoacílico ha llegado al sitio *A*. Las cinco bases "apiladas" del bucle del anticodón son ahora aquellas que están más cerca del extremo 3' de la

ro no el cuarto tipo, la uridina, que ocupa el lugar que tiene la timidina en el ADN. Además, ocasionalmente aparecen en el ARNt más de 50 versiones modificadas de las bases usuales.

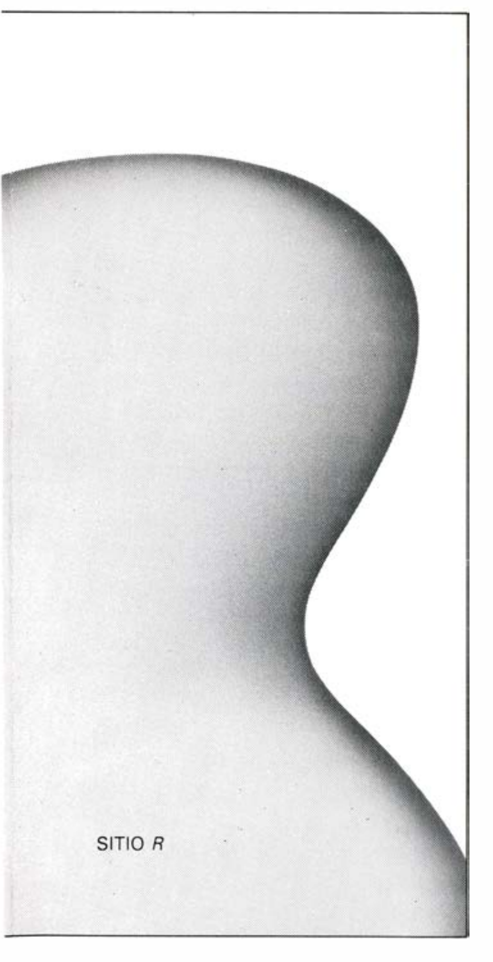
La estructura secundaria describe la forma en que ciertas secuencias de bases se unen (o forman pares de bases) con otras secuencias a lo largo de la cadena. Normalmente, la guanina se aparea con la citosina; son bases complementarias. De forma similar, la adenina se aparea con su complementaria, la uridina. En el ARN, alguno de los apareamientos no sigue este patrón. El punto importante, sin embargo, es que el ARNt puede formar cadena doble en los sitios donde la cadena sencilla se une consigo misma. Las regiones de cadena sencilla que quedan en medio forman bucles.

La estructura terciaria describe la forma en que la molécula de ARNt, con sus bucles y zonas de cadena doble, se dispone tridimensionalmente. En conjunto, un ARNt tiene forma de *L*. Al final del brazo corto de la *L* están los dos extremos libres de la cadena, uno muy cerca del otro. Uno de los extremos se denomina 5'. Se ha convenido en que sea el extremo por donde comience la numeración de bases de la secuencia primaria. El otro extremo, denominado 3', es el lugar por donde se une al ARNt un aminoácido específico. Al final del brazo largo de la *L* (y por tanto al otro extremo de la estructura terciaria de la molécula) se encuentra un bucle con siete bases. Las tres centrales (generalmente las números 34, 35 y 36) forman el anticodón: un triplete de bases que constituye el complemento del triplete que determina el aminoácido para el que este ARNt se ha especializado como transportador.

En el modelo de ciclo de elongación con dos sitios (*P* y *A*), los ARNt aminoacílicos llegan sucesivamente y por azar al sitio *A*. Cuando llega el correcto, las tres bases que forman su anticodón se unen en seguida a las tres bases del mensajero que forman el codón que va a leerse a continuación. La unión al ribosoma del ARNt correcto es, por tanto, más firme que la unión de uno incorrecto. Es lo que constituye el "reconocimiento" del ARNt correcto.

¿Hasta qué punto se acierta en el reconocimiento? Teniendo en cuenta la diferencia energética entre una unión correcta y una incorrecta entre codón y anticodón, se puede calcular que un ARNt incorrecto se unirá (y se incorporará un aminoácido incorrecto en la cadena naciente) aproximadamente una vez por cada 100 incorporaciones correctas. Las pruebas sugieren, sin embargo, que la tasa real de error es menor de una por cada 2000 incorporaciones. Por ejemplo, las proteínas más largas de *E. coli* son las subunidades β y β' de la polimerasa del ARN, cada una de las cuales tiene, aproximadamente, 1400 aminoácidos. Los errores en su secuencia son raros. Parece probable, por tanto, que el ribosoma debe mejorar la precisión inherente a la interacción entre codón y anticodón. Quizá sea ésta la función del sitio *R*. Quizá también esta mejora se realice a costa de un enlace fosfato rico en energía por cada aminoácido que se una a la cadena naciente.

¿Qué tipo de mecanismo molecular puede originar este aumento en la precisión? ¿En qué parte del ribosoma tiene lugar? ¿De qué forma puede estar involucrado el sitio *R*? Aunque no se puede dar todavía una respuesta definitiva, sí cabe aventurar algunas conjetu-





molécula. Este cambio permite al ARNt desplazarse sin producir "tirones" en el ARNm. Mediante un enlace entre bases de los dos bucles de los anticodones, puede asegurarse el alineamiento adecuado de los ARNt con el ARNm. La subunidad grande puede transferir ahora la cadena naciente al aminoácido del ARNt aminoacílico. El dibujo *5* completa el ciclo. Se expulsa el ARNt "desnudo"; el ARNt del sitio *A* ocupa el sitio *P* y el ARNm se reorganiza para la llegada del siguiente ARNt aminoacílico. No se han representado aquí ni el GTP ni el factor de elongación FE-G, que actúa en esta parte final.

ras. Después de todo, intentar encajar nuevos datos es uno de los placeres de la ciencia.

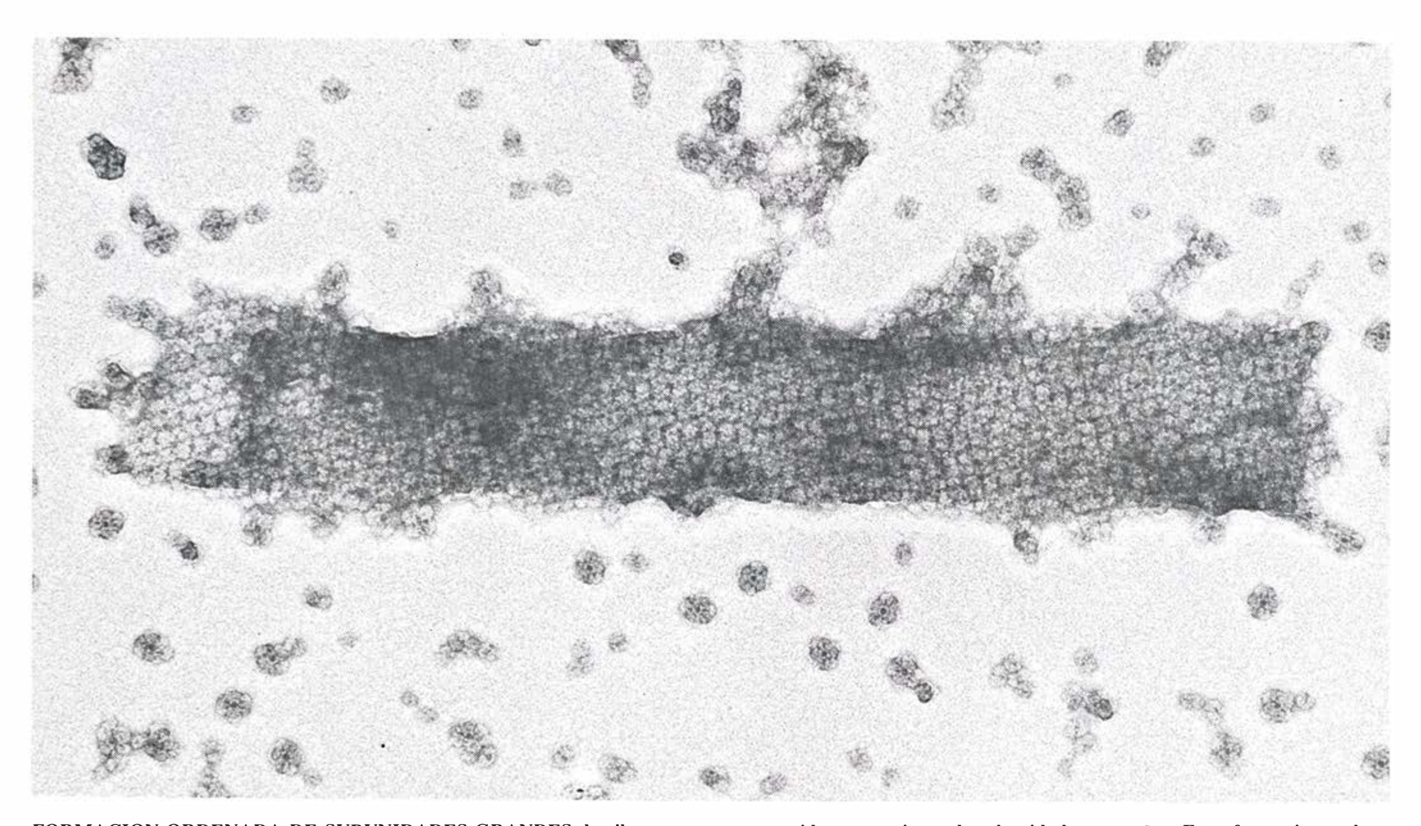
Requisitos del modelo

Cualquier mecanismo que mejore la precisión de la síntesis de proteínas debe ajustarse a una serie de observaciones básicas. Consideremos primero el sitio de unión del ARNm. En 1974, cuando se estaba desarrollando el modelo tridimensional de la subunidad pequeña, Kahan y el autor investigaban los resultados de varios experimentos de "soldadura". Grupos de la Universidad de Viena y del Instituto Max Planck de Genética Molecular, de Berlín, habían preparado análogos de ARNm en los que algunas bases habían sido sustituidas por otras que no se encuentran normalmente en el ARN. Cada análogo se une al ribosoma en la misma forma que lo haría un ARNm natural. Ahora bien, cuando el análogo se ha unido, un tratamiento adicional (generalmente irradiación con luz ultravioleta o adición de un agente químico "soldador") provoca una unión fuerte y permanente entre el análogo y aquello que le esté próximo en el ribosoma. La "soldadura" estableció la proximidad del ARNm a la proteína *S*11 y a la proteína vecina *S*18. La localización de aquellas proteínas en nuestro modelo nos permitió, a Kahan y a mí, proponer que la plataforma y la hendidura de la subunidad pequeña constituían el sitio más probable de unión para el ARN mensajero.

Un experimento realizado por Kahan con Robert Traut y John W. B. Hershey, de la Universidad de California en Davis, apoya esta conclusión. Consiguieron "soldar" el FI-3, un factor de iniciación que facilita la unión de los ARN mensajeros naturales, a tres proteínas. Las proteínas eran: *S*11, que nuestro modelo sitúa en la plataforma de la subunidad pequeña; *S*13, que nuestro modelo sitúa en la cabeza y *S*19, que aparece en la superficie de la subunidad pequeña en dos lugares de la cabeza. El significado de estos descubrimientos está claro. El anticodón de un ARNt debe unirse a un codón del mensajero. Por tanto, en cualquier mecanismo que se proponga, el anticodón de un ARNt unido al ribosoma deberá estar situado probablemente en la hendidura.

Consideremos a continuación los sitios de unión de los ARN de transferencia. Para empezar, de los experimentos de "soldadura" se puede deducir que el factor de iniciación FI-2 actúa en la hendidura y en la plataforma facilitando la unión del ARNt iniciador al ribosoma. Así, un modelo razonable colocará los sitios *A* y *P* en la hendidura. Una clave para situar el sitio *R* nos la da la localización de un grupo de proteínas –*S*3, *S*4, *S*5, *S*10, *S*12 y *S*14– en el lado de la subunidad pequeña que está más alejado de la subunidad grande. Las proteínas *S*4, *S*5 y *S*12 son las macromoléculas que los experimentos con estreptomicina y otras drogas implican en la fidelidad de la traducción del mensaje. Las otras tres proteínas –*S*3, *S*10 y *S*14– poseen una función distinta. Si alguna de ellas falta en un ribosoma que ha sido reconstituido en un tubo de ensayo a partir de una mezcla de los demás integrantes de la subunidad pequeña, el ribosoma resulta menos eficaz a la hora de unirse con los ARNt. Estos dos indicios implican a la parte externa de la subunidad pequeña en el reconocimiento de los ARNt aminoacílicos que llevan el anticodón correcto. Cabe presumir que el lugar de dicho reconocimento incluya el sitio *R*.

Obsérvese que todas las pistas anteriores implican a la subunidad pequeña en la tarea genética (o procesamiento de la información) del ribosoma: la unión de la cinta de instrucciones (es decir, el ARN mensajero) y la unión de sucesivos ARN de transferencia de acuerdo con aquellas instrucciones. Los experimentos de "soldadura" efectuados con un ARNt peptidílico en el que se ha modificado un aminoácido de la cadena naciente sugieren un papel para



FORMACION ORDENADA DE SUBUNIDADES GRANDES de ribosomas de *E. coli* conseguida por Michael Clark y el autor en la Universidad de California en Los Angeles, haciendo que las subunidades precipiten en una gota "colgante" de solución en una cámara con vapor de etanol. También se han conseguido agrupaciones de subunidades pequeñas. Estas formaciones abren la esperanza a que puedan someterse grandes cristales de ribosomas a difracción de rayos X, técnica que permite obtener un patrón de interferencia a partir de las regularidades atómicas de un cristal, y deducir así su estructura.

la subunidad grande. Cuando el ARNt peptidílico se une al ribosoma, un tratamiento adicional hace que el aminoácido modificado se una fuertemente a aquello que esté a su alrededor. De los resultados se deduce que la cadena naciente se sitúa cerca de la proteína *L*27, un constituyente de la subunidad grande, que ocupa la protuberancia central.

Ello no quiere decir que la proteína *L*27 sea la tan buscada transferasa del peptidilo: la responsable de que, en el ribosoma, los aminoácidos se unan para formar una cadena. Experimentos con subunidades grandes parcialmente ensambladas sugieren que la falta de cualquiera de las diversas proteínas que la forman ralentiza, pero no detiene, el alargamiento de la cadena. Si la transferasa del peptidilo existe, lo es como una propiedad de una región de la subunidad grande. En general, la subunidad pequeña realiza la tarea genética del ribosoma y la subunidad grande, la tarea bioquímica.

Un requisito final que debe cumplir un modelo aceptable se desprende de las investigaciones de Sigrid S. Thatch y Robert E. Thatch, de la Universidad de Washington en San Luis, y Peter Lengyel y sus colaboradores, de Yale. Los dos grupos utilizaron enzimas que degradan el ARN. Encontraron que la parte del ARN mensajero que se une al ribosoma (y está por tanto protegida de los enzimas) se desplazaba en una longitud correspondiente a tres bases, cuando el ARNt peptidílico se movía del sitio *A* al *P*. Esto es fácil de aceptar; no hay mayor inconveniente en imaginar un modelo donde la movilidad del ARNt del sitio *A* al *P* esté acoplada con el cambio de posición del mensaje, de forma que pueda leerse el siguiente codón. La exigencia más apremiante es que el ARN mensajero no se mueva durante el resto del ciclo de elongación. En el modelo que describiré a continuación, la interacción del anticodón del ARN de transferencia con el codón del ARN mensajero no se altera por el paso del ARNt del sitio *R* al *A*.

Mecanismo hipotético

Cada repetición del ciclo de elongación comienza (de acuerdo con el modelo) con el ARNt peptidílico en el sitio *P* y con los sitios *R* y *A* vacíos. Sucesivamente, los ARNt aminoacílicos van llegando al sitio *R*, que está situado en la parte externa de la subunidad pequeña. El sitio *R* se encuentra, por tanto, expuesto (muy apropiadamente) al citoplasma de la célula. Cada ARN de transferencia llega formando un complejo con FE-Tu y GTP, y la unión de este complejo al sitio *R* incluye la unión del anticodón del ARNt al codón que se va a leer a continuación. Hay un 99 por ciento de probabilidad de que se una al ARNt correcto. En cualquier caso, el ARNt que se une es desplazado (o empujado) inmediatamente desde el sitio *R* hasta el *A*, en la plataforma de la subunidad pequeña. La energía requerida para este desplazamiento procede de la molécula de GTP. Mientras el ARNt se mueve, permanece unido al ribosoma solamente por la unión codón-anticodón. La movilidad del ARNt constituye, pues, una prueba de la fuerza con que están apareadas las bases; es una segunda función del anticodón. Si la unión fuese más débil, el ARNt se "caería" del ribosoma.

Una analogía puede ayudar a comprender mejor el proceso. Suponga que tiene una pecera con peces plateados y dorados. Usted quiere coger los dorados. Para ello utiliza una red que introduce en la pecera, y cada vez que la saca obtiene una mezcla de peces en la que la proporción de dorados es mayor que la que existe en la pecera, pero todavía hay algunos plateados. Estos peces que va atrapando con la red los introduce a continuación en otra pecera, previamente vacía. Si usted ahora utiliza su red, con la misma habilidad, para coger los peces de la segunda pecera, la proporción de peces dorados que conseguirá, cada vez que saque la red, habrá mejorado. Lo que necesita en este experimento son dos peceras. En el ribosoma, las peceras son los sitios *R* y *A*.

Durante el desplazamiento de la molécula de ARNt desde el sitio *R* hasta el *A*, la interacción del codón con el anticodón no se altera. Para demostrar cómo es posible esto, debo antes explicar un nuevo aspecto de la estructura del bucle del anticodón de un ARNt. La disposición del bucle es el resultado de un compromiso entre dos requisitos. Por una parte, las siete bases que forman el bucle tienden a estar "apiladas" unas sobre otras. Mas, por tratarse de un bucle, la hilera de siete bases debe volver sobre sí misma. El compromiso a que se llega es que el bucle se ordena de tal suerte que sólo cinco de las bases están "apiladas". Si las cinco son las inmediatas al extremo 5' del ARNt, se dice que el bucle tiene una configuración 5'; si son las más cercanas al extremo 3', la configuración es 3'. De acuerdo con el modelo, el bucle del anticodón cambia su conformación de 5' a 3' cuando el ARNt se mueve desde el sitio *R* hacia el *A*. El ARNt puede, por tanto, cambiar su posición en el ribosoma sin provocar "tirones" en el mensaje.

Uno se imagina que cuando el ARNt aminoacílico llega al sitio *A* puede ya interaccionar con el ARNt peptidílico situado en el sitio *P*. Después de todo, el anticodón del ARNt aminoacílico que llega estará unido al codón inmediatamente posterior al traducido en el ciclo de elongación anterior: el codón al que permanece unido el ARNt peptidílico en esta etapa del ciclo. Si hubiese un "hueco" entre anticodones sucesivos (un error de "cambio de fase") significaría que un anticodón se aparearía con un codón incorrecto, y que por tanto se uniría al ribosoma un ARNt con un aminoácido equivocado. A partir de este error, la síntesis proseguirá totalmente equivocada: todos los aminoácidos serán incorrectos, a menos que nuevos errores de cambio de fase devuelvan, fortuitamente, la fase al mensaje.

El modelo permite la comprobación de la "fase de lectura" mediante apareamiento de bases entre ARNt adyacentes, situados en los sitios *A* y *P*. Un apareamiento entre las bases número 33 de ambos ARNt podría cumplir bien este papel. En primer lugar, la base número 33 está situada en el bucle del anticodón. En efecto, se encuentra próxima al anticodón; así, la unión entre ellas asegura que no haya errores de cambio de fase. Además, la estructura de un ARNt permite a la base número 33 "despegarse" de la molécula y participar en este enlace. Más aún, la base número 33 es constante en todos los ARNt que participan en el ciclo de elongación. Con independencia del aminoácido que transporte, la base 33 es uracilo. La única excepción es el ARNt especial que inicia la síntesis de proteínas en células eucariotas (células con núcleo). En este caso, la base número 33 es citosina. Esta diferencia puede tener su importancia. La naturaleza inusual del apareamiento de bases en el ARN determina que la citosina de un ARNt unido al ribosoma en el sitio *P* pueda aparearse con el uracilo de un ARNt unido al sitio *A*. Lo contrario no es el caso: el uracilo de un ARNt en el sitio *P* no puede aparearse con la citosina de un ARNt en el sitio *A*. Quizás eso impida que el ARNt iniciador entre en el sitio *A* durante el ciclo de elongación.

El apareamiento de uracilo con uracilo es, a pesar de todo, algo anecdótico. Lo importante del modelo estriba en la inclusión del sitio *R* y el desplazamiento del ARNt aminoacílico desde el sitio *R* hasta el *A*. Una vez el ARNt se haya unido al sitio *A*, prosigue el ciclo de elongación de acuerdo con el esquema aceptado. La cadena naciente se

transfiere al aminoácido del ARNt que está en el sitio *A*. Para ello no se requiere ninguna fuente de energía exterior. El FE-Tu, el GDP y el fosfato abandonan el ribosoma y llegan otras dos moléculas; a saber: el segundo factor de elongación, llamado FE-G, y una nueva molécula de GTP. El ARNt desacilado es expulsado del sitio *P*; el ARNt del sitio *A*, que ahora lleva la cadena naciente, gira en su sitio, y el mensaje avanza un codón, FE-G y la molécula de GTP (disociada ya, como su antecesora, en GDP y fosfato) abandonan el ribosoma, de forma que éste se encuentra listo para comenzar la adición programada del siguiente aminoácido.

ARN ribosómico

Del modelo se desprende un ribosoma grande, ya que debe albergar simultáneamente dos ARN de transferencia que interaccionan frente a frente. Esto es, el ribosoma debe poseer la magnitud suficiente para soportar un complejo molecular de 150 angstrom de longitud, aproximadamente, y al mismo tiempo ofrecer sitios de unión para el GTP y los diferentes factores de elongación.

¿Por qué el ribosoma está formado principalmente por ARN de gran tamaño? Quizá sean los residuos evolutivos de los ribosomas primordiales. F. H. C. Crick, del Instituto Salk de Estudios Biológicos, ha sugerido que, ya que las proteínas se hacen en los ribosomas, el primer ribosoma existiría antes que cualquier proteína ribosómica. Estaría formado por ARN. En favor de esta hipótesis abunda el hecho de que el ARNt está formado exclusivamente por ARN.

Algunos cálculos ofrecen una razón adicional. Supongamos que un orgánulo intracelular como el ribosoma requiere, en su estructura, una molécula cuyo peso molecular sea de un millón. Esa molécula podría ser un ácido nucleico: una cadena de nucleótidos. (Un nucleótido consta de una base, un azúcar y un grupo fosfato.) El peso molecular de un nucleótido es, aproximadamente, de 330; por tanto, se necesitarán 3000 nucleótidos. Las instrucciones genéticas para su secuencia estarían en una cadena de ADN de nucleótidos complementarios, y la "lectura" de las instrucciones para producir el ácido nucleico necesitaría 3000 pasos.

Pero valdría también una proteína: una cadena de unidades de aminoácidos. El peso molecular de un aminoácido es de 110, aproximadamente; por tanto, se necesitarían unas 9000 unidades. Las instrucciones genéticas para su secuencia estarían en una cadena de ADN, con un codón, o tres nucleótidos, por cada aminoácido. La lectura de las instrucciones para producir el ARN mensajero que codifique la proteína requeriría 27.000 pasos. Supongamos que la traducción del mensaje por un ribosoma está libre de error, y que sólo la transcripción del ADN en ARN es imperfecta. Aún así, es evidente que, para cualquier probabilidad de error dada, se puede sintetizar ARN con un peso nueve veces mayor que el de una proteína.

Imaginemos, por último, que un error en la transcripción del ADN a ARN sitúa un nucleótido extra en algún lugar del ARN. Esta adición quizá sólo produzca una ligera comba en la estructura terciaria del ARN. Pero si el ácido nucleico fuera un ARN mensajero, el nucleótido extra ocasionaría un cambio de fase, y la proteína sintetizada a partir de las instrucciones del mensajero probablemente no sería útil. De esta forma, si para la estructura se necesita una molécula muy grande, la precisión en la síntesis de la molécula probablemente determine que, además, sea un ARN.

Los progresos registrados, a lo largo de los últimos años, en el conocimiento de la estructura y función del ribosoma aumentan las esperanzas de conseguir mayores avances. En ese sentido, las indicaciones de que los ribosomas poseen sólo una única estructura y configuración abren nuevas esperanzas de que puedan cristalizarse para los análisis de difracción de rayos X. En efecto, Michael Clark, Jerome Langer y el autor acaban de conseguir recientemente formaciones ordenadas de subunidades ribosómicas de *E. coli* utilizando la técnica de la "gota colgante". Las subunidades (en una gota de solución) penden de una lámina de cristal, en una cámara que contiene vapor de etanol. Cuando la concentración de etanol en la gota se aproxima al 10 por ciento, las subunidades comienzan a precipitar.

Quizás, a más largo plazo, puedan conseguirse formaciones mayores y más ordenadas de las alcanzadas hasta ahora, para someterlas a análisis de difracción de rayos X. Con todo, una visión del ribosoma con la localización de cada átomo no mostraría necesariamente qué átomo pertenece a cada molécula. Las técnicas que utilizan anticuerpos y el microscopio electrónico seguirán siendo, sin duda, inapreciables. Todos los esfuerzos tenderán a un mismo objetivo: descifrar, en todos sus pormenores, de qué manera los ribosomas fabrican las proteínas.

La forja de los metales

Este antiguo arte conforma los metales mediante golpe o presión. La forja actual está evolucionando merced a un mayor conocimiento de la deformación y el uso del ordenador en distintas etapas del proceso

S. L. Semiatin y G. D. Lahoti

En sus inicios, se aplicaba la forja para dar forma a una pieza de metal sujetándola sobre un yunque y golpeándola con un martillo. Cabe presumir que lo que el hombre de la antigüedad pretendía era obtener formas útiles. Por forja se entiende hoy la deformación por compresión de un material colocado entre matrices. Estas, a menudo, son componentes de prensas de gran tamaño capaces de ejercer una fuerza enorme, y la pieza obtenida puede ser hasta la compleja estructura de las alas de un avión. Aunque se podría suponer que una tecnología tan antigua apenas ofrecería posibilidades de perfeccionamiento, la verdad es que las técnicas de forja han variado mucho merced a los nuevos conocimientos sobre las propiedades de deformación de los metales y en virtud de la creciente aplicación de los ordenadores al diseño de matrices y otros equipos de forja y al análisis del comportamiento de una pieza de metal en el proceso de forja.

La forja moderna ocupa un lugar prominente en lo que se denomina hechurado primario, subsiguiente a las etapas de elaboración metálica, a lo largo de las cuales el metal se extrae de la mena, se funde para eliminar impurezas o para añadir elementos aleantes y luego se moldea. A veces, ésta es su forma final; otras, se obtiene un lingote que sufre una deformación posterior. La forja comprende casi la mitad del hechurado primario, que incluye también procesos como el extrusionado, el estirado y el laminado. Suele suceder que una forma hecha por uno de esos procesos se someta luego a un proceso de hechurado secundario, como el estampado o el mecanizado.

Los documentos más antiguos, que describen el simple martilleo del oro y del cobre en varias regiones del Oriente Medio, se remontan unos 8000 años a. de C. En comparación con los hechurados modernos, el de estos metales resultaba muy deficiente, pues se desconocía el arte del afino por fusión con reacción y la capacidad de forjar el metal estaba limitada por las impurezas que quedaban después de la obtención del mineral a partir de la mena. La aplicación de la fusión con reacción del cobre, hacia el 4000 a. de C., supuso un método útil de purificación de los metales mediante reacciones químicas en estado líquido. Más tarde, en la Edad del Cobre, se estableció que el martilleo del metal provocaba un aumento de su resistencia (debido a lo que se denomina hoy endurecimiento por deformación en frío del metal). La necesidad de utilizar materiales de mayor resistencia a los golpes condujo al descubrimiento de aleaciones muy resistentes a base de cobre y estaño (Edad del Bronce) y de hierro y carbono (Edad del Hierro).

La Edad del Hierro fecha su punto de arranque en torno al año 1200 a. de C., trece siglos después de los comienzos de la Edad del Bronce. La razón de esta demora hay que buscarla en la ausencia de un método por el que conseguir las altas temperaturas necesarias para fundir y afinar la mena de hierro. Más tarde se descubrió que las aleaciones férreas se forjaban fácilmente a temperaturas superiores al sesenta por ciento de la temperatura de fusión absoluta expresada en grados Kelvin, o grados Celsius por encima del cero absoluto. La forja a tales temperaturas se llama ahora forja en caliente o hechurado en caliente. (Uno de sus resultados, además de la forma del metal, era la aparición de óxidos de hierro en forma de cascarilla negra. La presencia constante de este material fue lo que en lo primitivos talleres de forja dio nombre inglés al operario: "blacksmith", forjador.)

El desarrollo del martillo de palanca, en el siglo XIII, alivió grandemente el trabajo de la forja. Esta máquina empleaba energía hidráulica para elevar un mango que contenía un martillo en su extremo; se llamó también martillo inclinado porque el brazo estaba inclinado. Después de levantar el martillo se le dejaba caer para que, así, generase el golpe de la forja. Ese aparato, relativamente simple, permaneció en servicio durante varios siglos, aplicado sobre todo a la forja de barras y láminas. Con la revolución industrial aparecieron procesos para la fabricación de hierro y acero en grandes cantidades, junto con un gran aumento de la demanda de productos metálicos, lo que requería sistemas de forja de mayor capacidad. Esta exigencia se satisfizo con el desarrollo del martillo de alta velocidad impulsado por vapor, donde el martillo se elevaba mediante la energía del vapor, y con el desarrollo de una prensa hidráulica, donde la fuerza se suministraba por presión hidráulica. De estas máquinas de forja salieron una gama de productos que iban desde armas de fuego hasta piezas de locomotoras.

En los últimos cien años se han desarrollado nuevos tipos de equipos de forja y nuevos materiales con propiedades y aplicaciones especiales. El aprovechamiento de esos materiales ha sido posible gracias al conocimiento (adquirido sobre todo en los últimos veinticinco años) de los efectos de la temperatura y de la velocidad de deformación en las características básicas del hechurado de los metales, de los efectos del diseño y de la lubricación de las matrices en la fluencia de los metales y, en tercer lugar, de la fuerza necesaria para la forja. La aplicación de este conocimiento al trabajo de forja se ha visto facilitada por el desarrollo de refinados análisis matemáticos de los procesos de forja, análisis que se realizan con ordenadores digitales de alta velocidad.

Se acepta comúnmente que la capacidad para deformar metales se funda en la naturaleza del enlace metálico. En este tipo de enlace, los átomos del metal comparten sus electrones externos. El enlace es también adireccional, de modo que los átomos pueden situarse muy próximos unos a otros en disposiciones cristalinas regulares tridimen-

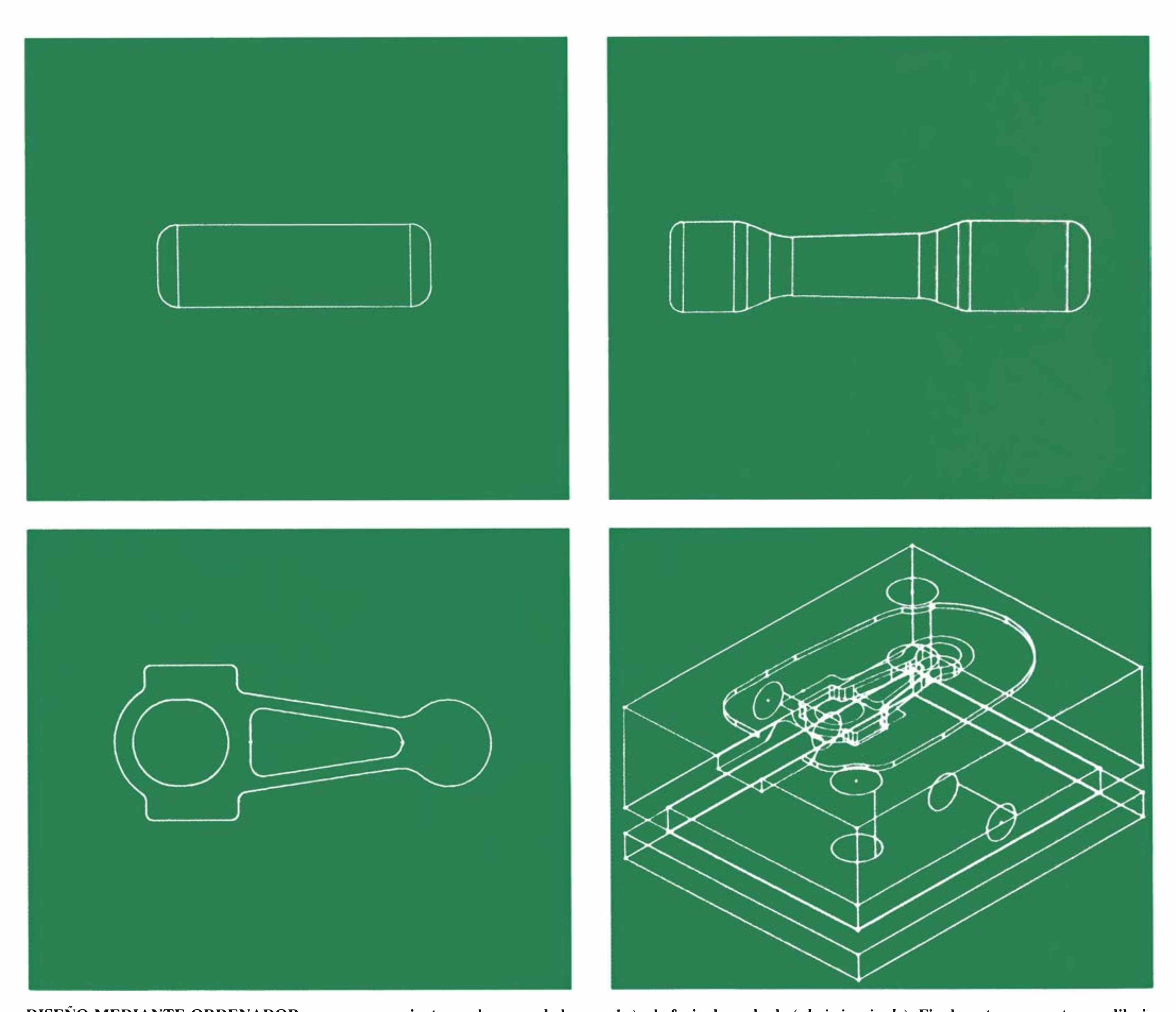
sionales. La presencia en estas estructuras de dislocaciones, o imperfecciones en la disposición tridimensional, posibilita que planos compactos de la red cristalina se deslicen a lo largo de direcciones de máximo empaquetamiento. Como se puede suponer, la deformabilidad relativa de los diferentes metales depende, en alto grado, de la facilidad con que pueden generarse las dislocaciones para deformar el metal sin causar su rotura o sin introducir en él otros tipos de defectos.

La facilidad con que las dislocaciones favorecen la plasticidad, sin resultados indeseables, es función de las características estructurales, de la temperatura del metal y de la velocidad de deformación. Entre las características estructurales cabe citar la disposición cristalina tridimensional, la presencia de elementos aleantes y la existencia de otras dislocaciones, aunque la característica más importante sea, tal vez, el tamaño de grano. Por granos se entiende los bloques de construcción de los agregados metálicos. Cada grano es un cristal distinto, con una orientación de la red cristalina diferente de la de sus vecinos. Cuando el tamaño de grano es grande en relación con la magnitud global de la pieza de trabajo (la pieza de metal que se está hechurando), se pueden producir grietas y propagarse fácilmente a lo largo de los límites de grano.

Este problema se agudiza cuando la estructura del metal es típica de una pieza moldeada, por ejemplo una pieza que ha solidificado desde el estado líquido en un molde. En primer lugar, solidifica el metal en contacto con las paredes frías del molde. Crecen grandes granos en forma de columnas de varios centímetros. Resulta difícil forjar entonces el metal sin su rotura, porque los planos entre las caras de los granos son frágiles.

Por ese motivo la forja de metales moldeados suele desarrollarse en dos etapas. Sirve la primera para destruir la estructura de granos grandes, dando lugar a una estructura de forja químicamente homogénea con un tamaño de grano mucho menor, normalmente medible en centésimas de milímetro, y de superior deformabilidad. Con metales de esta estructura se fabrican los componentes forjados acabados.

Aunque la mayoría de los metales forjados pueden trabajarse en una amplia gama de temperaturas y velocidades de deformación, debe tenerse cuidado, sin embargo, al seleccionar la temperatura de forja. La forja en caliente requiere el mínimo de energía en el forjado y produce el máximo de de-



DISEÑO MEDIANTE ORDENADOR, un avance reciente en el campo de la forja de los metales. La pantalla de un tubo de rayos catódicos conectada a un diseñador muestra un conjunto de barras de partida para una biela de un motor de automóvil (*arriba izquierda*), una forja intermedia (*arriba derecha*) y la forja de acabado (*abajo izquierda*). Finalmente, se muestra un dibujo de una de las dos matrices empleadas en la operación final de forja. Las técnicas de diseño en forja mediante ordenador han sido desarrolladas por A. A. Badawy y C. F. Billhardt, del Battelle Memorial Institute de Columbus.

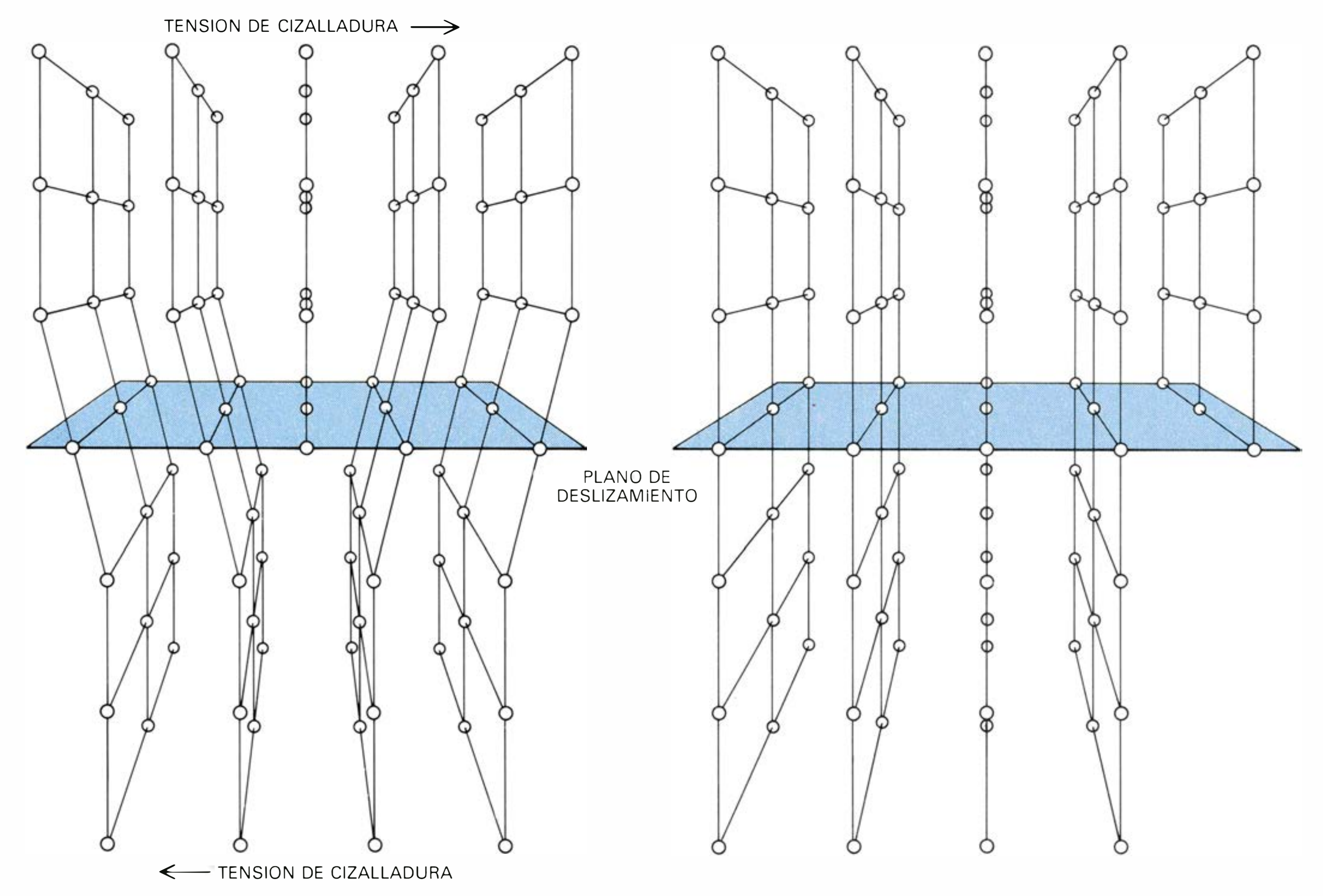
formación, pero al forjador le cuesta controlar las dimensiones del producto porque el metal no se contrae uniformemente cuando se enfría. De ahí que la forja se realice a menudo a temperatura ambiente, a la que se halla el régimen de trabajo en frío de la mayoría de los metales (25 por ciento o menos de la temperatura absoluta de fusión del metal). A tales temperaturas, la deformabilidad del metal es algo menor que las temperaturas de trabajo en caliente, pero es mucho mayor que en régimen de trabajo a temperatura media (entre el 25 y el 60 por ciento de la temperatura absoluta de fusión). Volveremos luego sobre esta aparente paradoja.

Ni que decir tiene que el principal objetivo de la forja es hechurar partes de tamaño correcto. Sentada esa premisa, los puntos principales a considerar son la fuerza necesaria para conseguir la deformación (en otras palabras, la capacidad del equipo) y la deformabilidad del metal. De rango secundario, principalmente de índole económica, son la tasa de producción y la duración de las matrices.

¿Cómo se calcula la fuerza necesaria para la deformación? Esa fuerza, que se denomina carga de forja, se deduce, para un metal particular, de la compresión de muestras cilíndricas entre matrices planas y paralelas que han sido bien lubricadas. Este modo de deformación simula el tipo más sencillo de operación de forja. Por medio de este ensayo, la resistencia a la plasticidad, o resistencia a la deformación, en términos de tensiones (carga dividida por área), se representa como una función de algún dato de la deformación, normalmente la reducción de la muestra (expresada como el logaritmo del cociente entre la altura inicial y la altura final) o bien la disminución de altura expresada en porcentaje. Las curvas tensión-deformación, así obtenidas, muestran el efecto del grado de deformación, de la velocidad de deformación y de la temperatura sobre la resistencia del metal a la deformación. Estas gráficas son esenciales para diseñar un modelo matemático del proceso de la forja.

A temperaturas de trabajo en frío las tensiones son elevadas. Ello se debe, principalmente, a la dificultad de mover las dislocaciones a través de la red y a que éstas se multiplican rápidamente, conduciendo a un endurecimiento por trabajo: aumento de la tensión con la deformación. Los límites de deformabilidad son modestos. A temperaturas de trabajo medias, los niveles de tensión son algo menores, pero existe todavía una gran cantidad de endurecimiento por trabajo y la deformabilidad es baja.

A temperaturas de trabajo en caliente, los niveles de tensión bajan todavía más, porque la energía térmica ocasiona grandes fluctuaciones de los átomos en la redes cristalinas de los granos del metal alrededor de sus posiciones de equilibrio; de ahí que resulte más fácil el desplazamiento de las dislocaciones a través de las redes. Además, las curvas tensión-deformación no reflejan endurecimiento por trabajo, ya que parte de la energía térmica se consume en "recocer" muchas de las dislocaciones generadas durante la deformación. La estructura de un metal puede cambiar tan drásticamente durante el trabajo en caliente que el resultado es un ablandamiento (una disminución de la tensión con la deformación). Los mecanismos



DEFORMACION DE UN METAL originada por el movimiento de dislocaciones a través de su red cristalina. La dislocación constituye un defecto en la disposición regular de los átomos en el cristal. Una "dislocación en arista" se genera por la presencia de un semiplano extra de átomos en una red regular (*izquierda*). La dislocación facilita el deplazamiento de la red en un "plano de deslizamiento", que normalmente es un plano de empaquetamiento compacto, y a lo largo de direcciones de empaquetamiento compacto (*derecha*). Cuando la presión de forja se aplica a la pieza de metal, cada cristal del metal se deforma a lo largo de varios conjuntos de planos y direcciones, con objeto de acomodar la deformación impuesta a la pieza por la operación de forja.

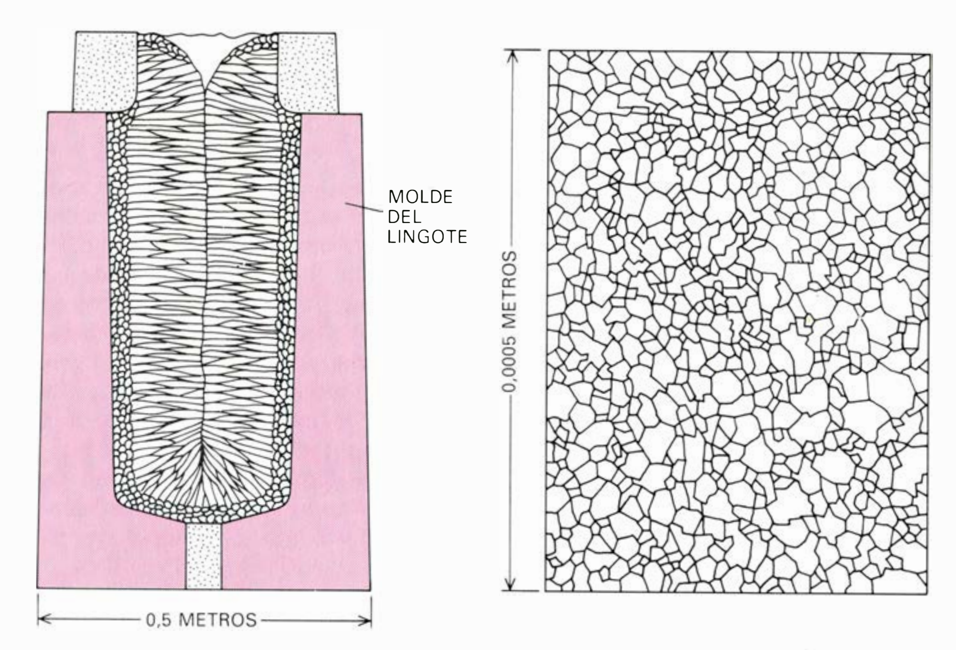
microscópicos de este cambio han sido documentados por John J. Jonas, de la Universidad de McGill, y por Michael Luton, de la Exxon Corporation, que han demostrado que el ablandamiento puede ocasionar una exagerada deformación no uniforme en muestras de compresión.

En una operación de forja real, la carga y la presión de forja dependen marcadamente de la fricción entre las matrices y la pieza de trabajo, así como del límite de fluencia del metal de la pieza. Bajo condiciones de lubricación perfecta, la presión requerida para forjar una muestra cilíndrica es uniforme e igual al límite de fluencia del material. Con la fricción, la distribución de la presión de forja adquiere la forma de una tienda de campaña: la presión es mínima (igual al límite de fluencia) en la parte exterior de la muestra y máxima en el centro. Asociadas con esta distribución de la presión están las tensiones de compresión radial en la muestra, las cuales se desarrollan a partir de los efectos de constricción por la fricción. Dependiendo de la magnitud de la fricción y de la deformación, la presión máxima en la forja puede tomar un valor diez veces superior al límite de fluencia del material.

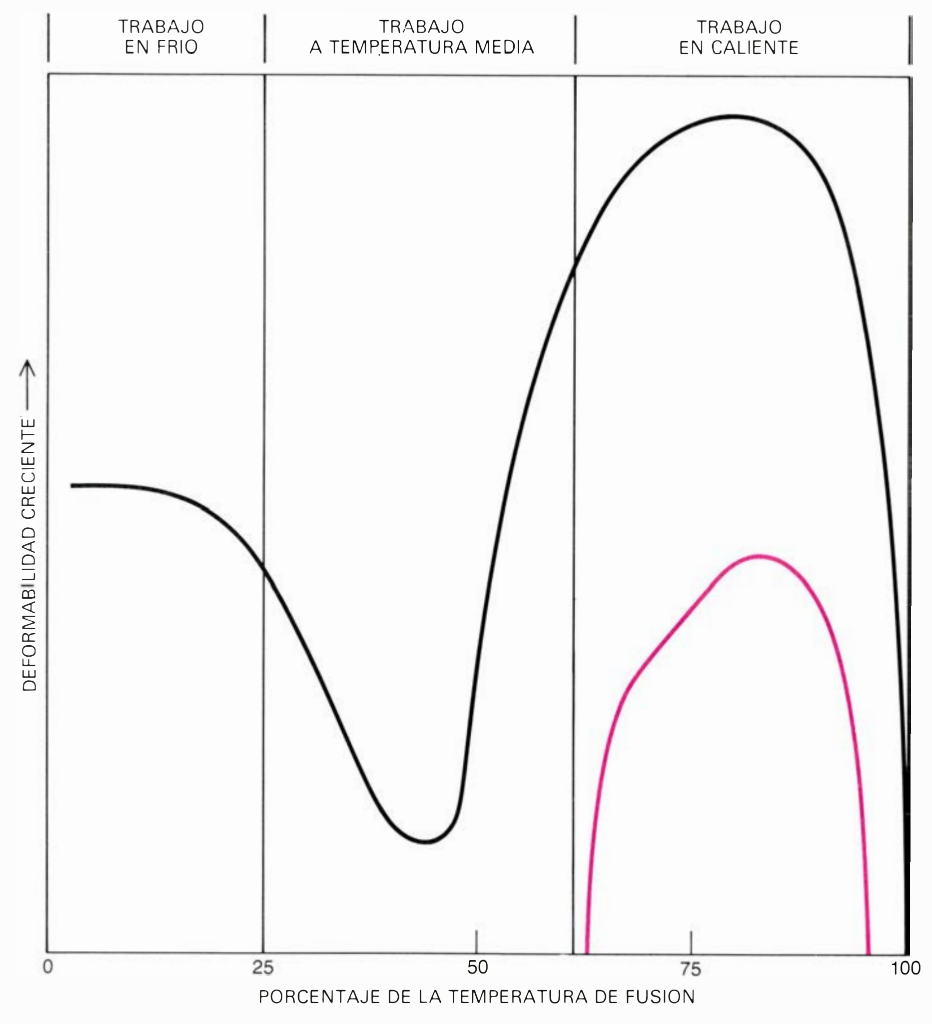
Las condiciones de fricción en la interfaz de forja están gobernadas por la lubricación y, en forja a temperatura media y en forja en caliente, por el gradiente de temperatura entre las matrices frías y el material caliente. (Este gradiente puede conducir a un enfriamiento del material y a un "gripado" del material que toca a las matrices.) Los lubricantes reducen la fricción y controlan o mejoran la uniformidad de fluencia en el metal. También reducen el desgaste de la matriz y evitan que la pieza de trabajo se pegue a las matrices.

En la forja en frío de los aceros, suelen usarse jabones como lubricantes. Pues no sólo proporcionan una lubricación excelente a temperatura ambiente, sino que pueden eliminarse fácilmente también de las partes forjadas, sin afectar la superficie del componente. En la forja a temperatura media y en caliente, el lubricante más común es probablemente el grafito. Los grafitos con aceite tienden a expulsar humos, y por eso la última tendencia en la industria de forja apunta hacia el uso de grafitos a base de agua.

Los lubricantes para forja se controlan por medio del "ensayo del anillo", que consiste en deformar una muestra anular, de dimensiones conocidas, entre dos matrices planas y lubricadas



LA DEFORMABILIDAD DE UN METAL depende del tamaño de los granos (cristales individuales) en la pieza metálica que va a ser forjada. Un metal que ha sido moldeado en un lingote (*izquierda*) presenta granos grandes, medibles en centímetros, y normalmente es difícil de forjar. Una estructura de forja (*derecha*) tiene granos pequeños, medibles en centésimas de milímetro, y de forja bastante más fácil. Es por esto por lo que la forja suele realizarse en dos etapas. En la primera de ellas, se aplica la forma para destruir los granos grandes y conseguir la estructura apropiada para realizar la forja de acabado.



INFLUENCIA de la temperatura en la deformabilidad de los metales, cuando se trata de estructuras moldeadas (*color*), que tienen granos gruesos, y en el caso de estructuras de forja de grano fino (*negro*).

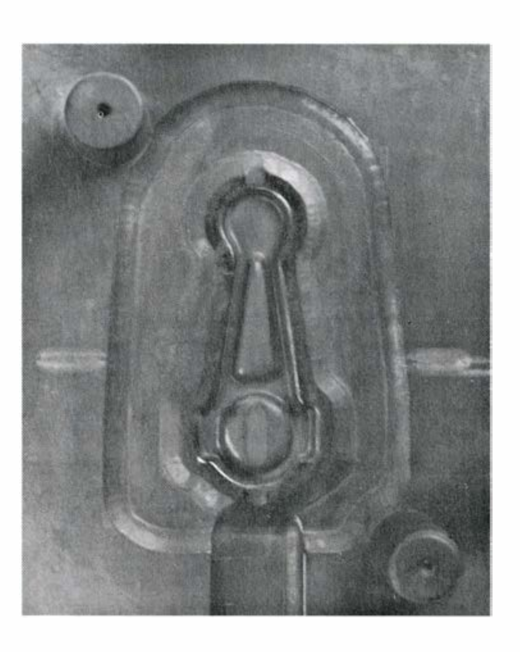
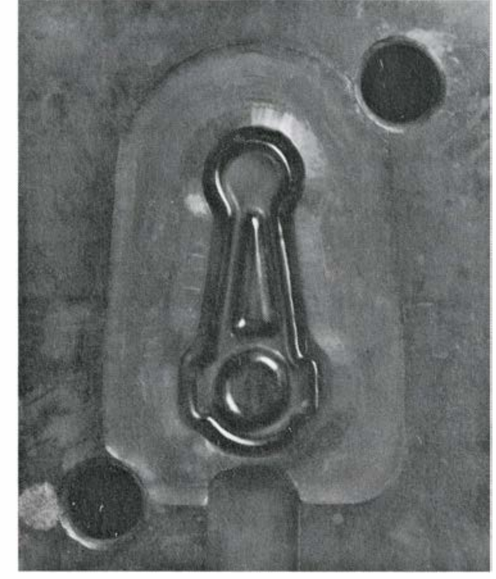
hasta una reducción predeterminada de su altura. El cambio en el diámetro interior de la muestra anular es muy sensible a la fricción; la magnitud del cambio aumenta si la fricción es pequeña y decrece, si se eleva. El porcentaje de cambio del diámetro interior proporciona un criterio para comparar la eficacia de los lubricantes.

Cuando la parte que se va a tratar tiene una forma sencilla, resulta relativamente fácil estimar la carga de forja mediante las mediciones de datos sobre la resistencia a la fluencia y a la fricción. Sin embargo, las estimaciones de la carga son difíciles de obtener para una pieza de forma complicada. Los cálculos dependen, principalmente, de la preforma, o forma de partida, de la pieza de trabajo, del diseño de la matriz y del grado de contacto entre la matriz y el material. El grado de contacto es pequeño en la forja de matriz abierta y, elevado, en la forja de matriz cerrada.

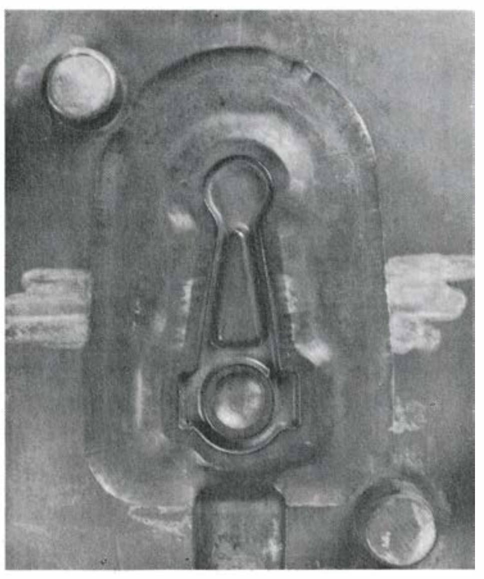
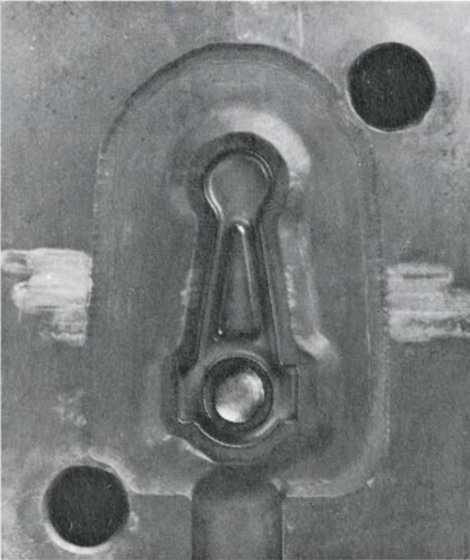
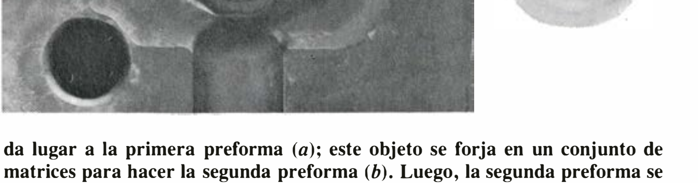
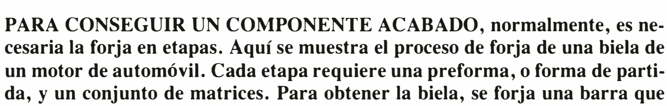
Además, la manufactura de componentes típicamente complicados requiere varias etapas de forja. El diseño del número y configuración de las formas intermedias (conocidas como preformas), las matrices y las condiciones de forja para las diferentes etapas constituye, a buen seguro, la tarea más difícil del ingeniero. La forja de una biela para un motor de automóvil puede precisar hasta tres etapas antes de obtener la forma acabada. En los talleres, los ingenieros de diseño experimentados suelen seguir líneas de trabajo empíricas para llegar a sus decisiones. Ensayan materiales como la cera, el plomo y la arcilla en matrices a pequeña escala o examinan la dirección del flujo estudiando los efectos de deformación en enrejados previamente trazados sobre las probetas. El análisis de estos enrejados sirve también para descubrir regiones donde puedan desarrollarse condiciones de tensión indeseables en el curso de la forja. Así, esta técnica revela regiones donde efectos geométricos pueden conducir a esfuerzos de tracción, denominados esfuerzos de tracción secundarios, incluso bajo la llamada carga compresiva. En la actualidad, se están aplicando técnicas de ordenador para el diseño de preformas y matrices y simular la fluencia de los metales.

Una vez diseñadas las matrices y determinada la carga de forja, se selecciona el material de la matriz y del equipo



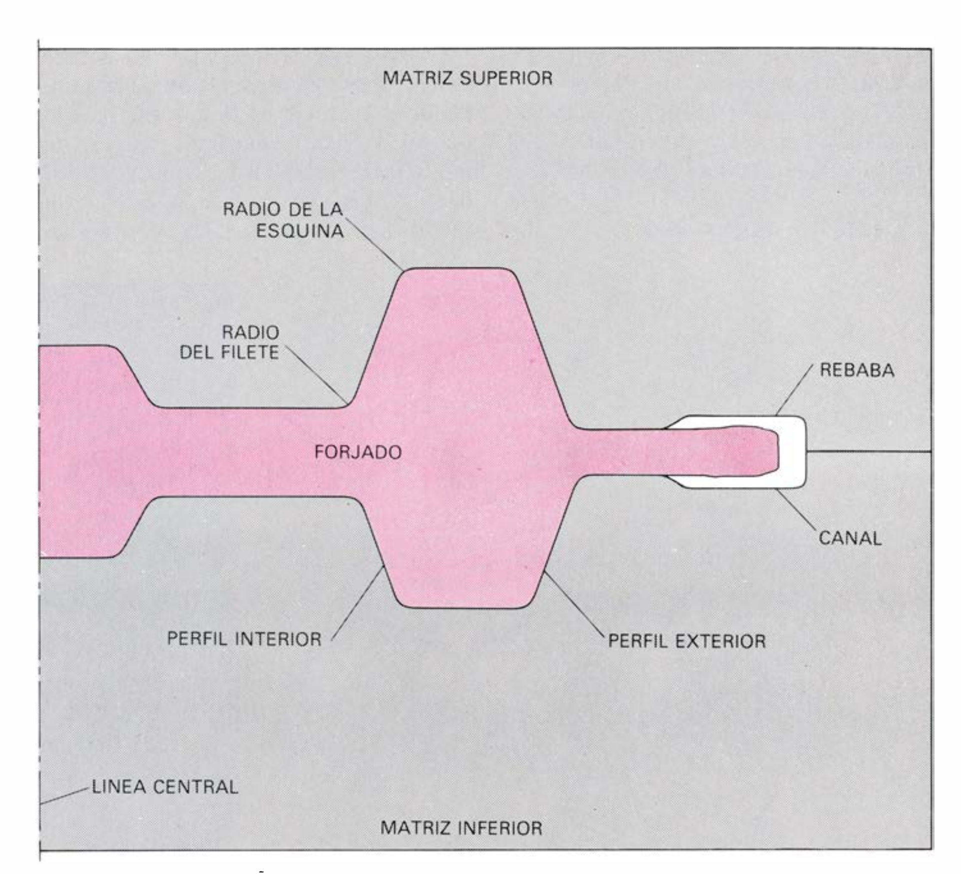





PARA CONSEGUIR UN COMPONENTE ACABADO, normalmente, es necesaria la forja en etapas. Aquí se muestra el proceso de forja de una biela de un motor de automóvil. Cada etapa requiere una preforma, o forma de partida, y un conjunto de matrices. Para obtener la biela, se forja una barra que da lugar a la primera preforma (*a*); este objeto se forja en un conjunto de matrices para hacer la segunda preforma (*b*). Luego, la segunda preforma se forja con otro conjunto de matrices que originarán la biela (*c*); y, a ésta, se le quitarán las rebabas (metal en exceso) para adoptar la forma de acabado.

para la forja. Los materiales para las matrices pueden ser desde aleaciones corrientes y aceros para herramientas de alta resistencia hasta las superaleaciones que pueden soportar altas presiones a elevadas temperaturas. Cuando la presión y la temperatura de forja son moderadas, y breve el tiempo de contacto, se muestran idóneas las matrices de acero aleado. A veces, una matriz de este tipo se recubre con una superaleación dura, resistente al desgaste, al objeto de aumentar la vida de la matriz a un coste razonable. El mismo fin se consigue, en muchas operaciones de forja en frío, por inserción de carburo de wolframio en la matriz: pequeñas matrices en un soporte. En operaciones exóticas de forja en caliente, en que vienen a coincidir las temperaturas de la matriz y de la pieza, el conjunto de la matriz se hace con materiales especiales (tales como IN-100, una aleación de níquel, o TZM-molibdeno) y se la calienta hasta 1000 grados Celsius en una atmósfera controlada para evitar la oxidación. Este tipo de operación de forja se emplea en el hechurado de precisión de aleaciones de determinados metales: titanio, níquel y berilio.



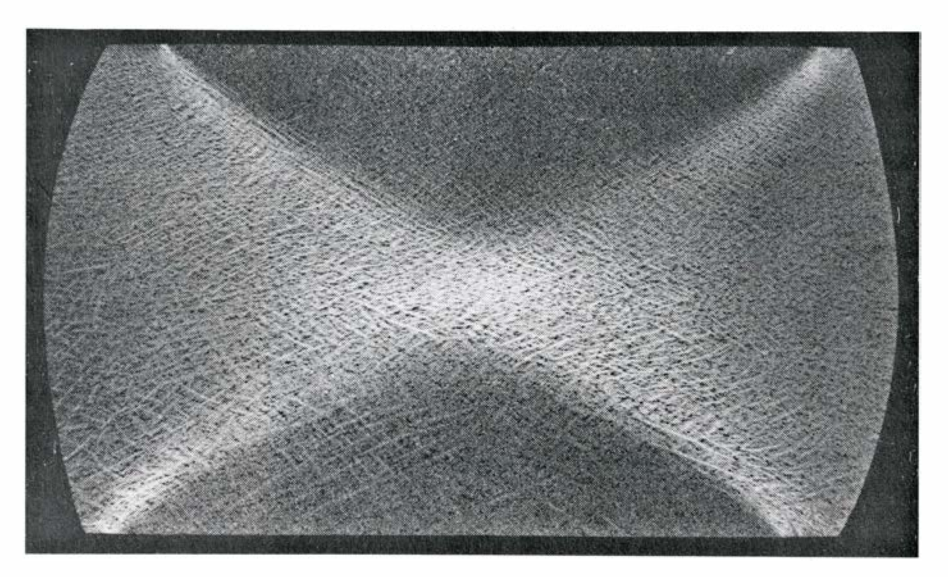
ELEMENTOS DE DISEÑO para matrices de forja, mostrados en una sección transversal vertical a través de las mitades derechas de un par de matrices y la mitad derecha de una pieza forjada. El diseño de las matrices requiere la determinación de los radios del filete y de la esquina, los cuales aumentan la fluencia del metal; los ángulos de los perfiles interior y exterior, que facilitan la eliminación de la pieza una vez ha sido forjada, y los canales verosímiles hacia los que puede moverse el material sobrante.

Por último, el ingeniero selecciona el equipo más apropiado para forjar una determinada pieza. Atiende primero a la fuerza y energía que la máquina de forja puede ejercer, pero también se toman en consideración otras características, la velocidad de la máquina por ejemplo, que pueden afectar a la deformabilidad de la pieza.

A la hora de la selección, el ingeniero contempla tres tipos básicos de equipos de forja: el martillo pilón (el mango puede caer por mera gravedad o con auxilio de energía), la prensa mecánica y la prensa hidráulica. Con el martillo de forja se dan golpes sucesivos a la pieza de trabajo; mientras que en la prensa mecánica o hidráulica la deformación suele realizarse en una sola operación. Desde otro punto de vista, las diferencias principales que existen entre los tres tipos de equipos residen en su velocidad de forja y en el modo de almacenar energía. Los martillos son los más rápidos, con velocidades que van de tres a nueve metros por segundo. La energía almacenada por éstos es energía potencial, debida al peso elevado, que se convierte en cinética al caer el mango. En un martillo auxiliado con energía se aplica vapor o aire a presión al mango durante la caída. Los martillos se clasifican de acuerdo con la energía que pueden liberar; los martillos grandes tienen capacidades de unos 680.000 newton-metro. Los martillos son las máquinas de forja menos caras y más versátiles.

La energía de una prensa mecánica se almacena en una rueda de inercia. El mango se acciona mediante un sistema mecánico: una manivela o una leva. Ya que una prensa mecánica tiene un golpe que puede ser controlado con precisión, se la elige a menudo cuando hay que forjar con gran control dimensional. Una prensa mecánica grande puede aplicar cargas que superan los 90 meganewton (10.000 toneladas).

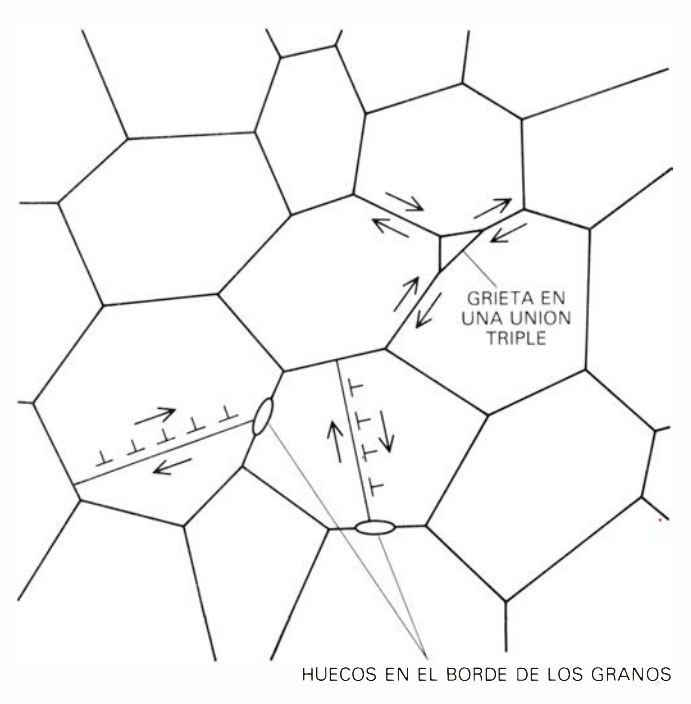
Las prensas hidráulicas son las más lentas (0,06 a 0,30 metros por segun-



LAS BANDAS DE DESLIZAMIENTO definen un tipo de defecto que puede aparecer en la forja en frío. Vienen determinadas por las regiones más claras de la sección transversal de un cilindro de acero al carbono que fue comprimido entre matrices no lubricadas. Las bandas aparecen cuando la fricción constriñe la parte del metal deformado en contacto con las matrices. Estas áreas se vuelven zonas de metal "muerto"; son las regiones más oscuras. Las zonas de metal muerto están separadas de las restantes zonas del metal deformado por regiones (bandas) donde el grado de deformación es mayor que el promedio.

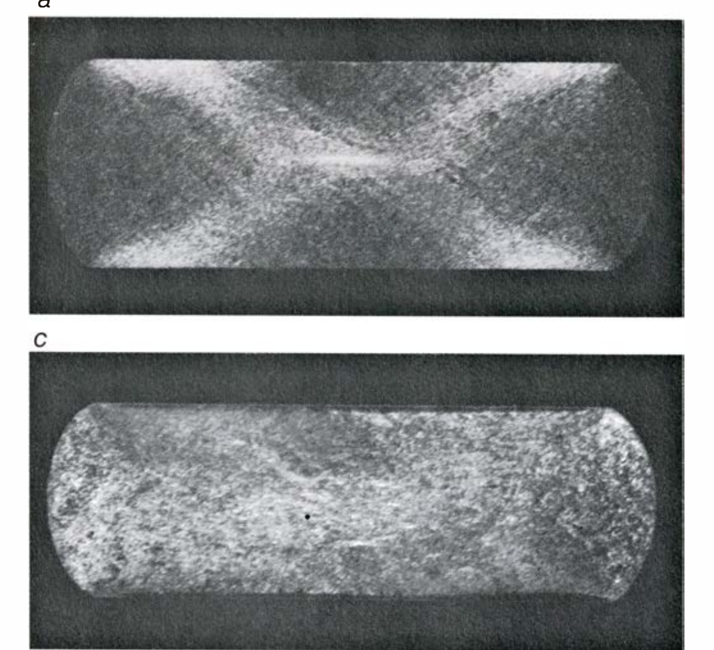
do), porque son accionadas mediante un gran pistón guiado por un fluido sometido a elevada presión. Algunas de las prensas de forja más grandes del mundo son hidráulicas. Las dos mayores de los Estados Unidos desarrollan cargas de 445 meganewton.

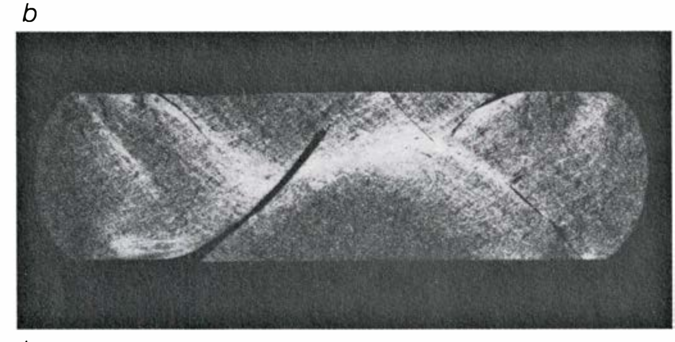
La segunda consideración en importancia para la determinación de las condiciones de forja es la deformabilidad del metal. Como se puede esperar, el modo de rotura de los metales forjados depende grandemente de la temperatura. En la forja en frío, las roturas suelen desarrollarse como grietas en las superficies que no están en contacto con las matrices. La rotura resulta de un abultamiento o pandeo debido a la fricción de aquellas superficies, los cuales, a su vez, conducen a la formación de esfuerzos de tracción secundarios. Pe-

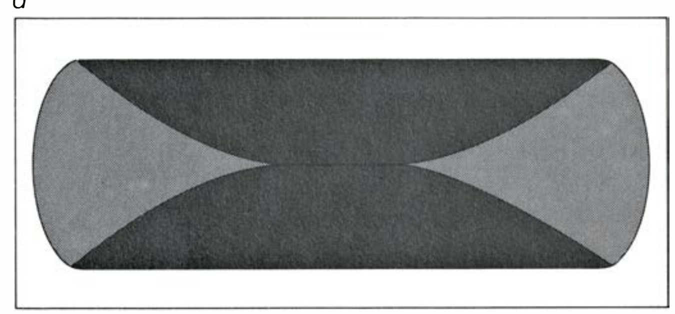




DOS PROBLEMAS que limitan el hechurado del metal a temperatura media y a temperaturas de trabajo en caliente son la formación de grietas en las uniones triples y cavidades que aparecen en los límites de grano. Estos problemas se esquematizan en la ilustración de la izquierda y se muestran en una microfotografía de una aleación de titanio, que aparece a la derecha. En el diagrama esquemático, las flechas denotan esfuerzos de cizalladura y los símbolos de forma de *T* invertida señalan dislocaciones. Las grietas en las uniones triples resultan de una cizalladura activada térmicamente en los límites de grano. Las cavidades se inician por las acciones de las dislocaciones que impulsan a las partículas hacia los límites de grano. En la microfotografía, ampliada unos 250 diámetros, se distinguen los límites de los tres granos de titanio como finas cintas blancas que convergen hacia la grieta de la unión triple.







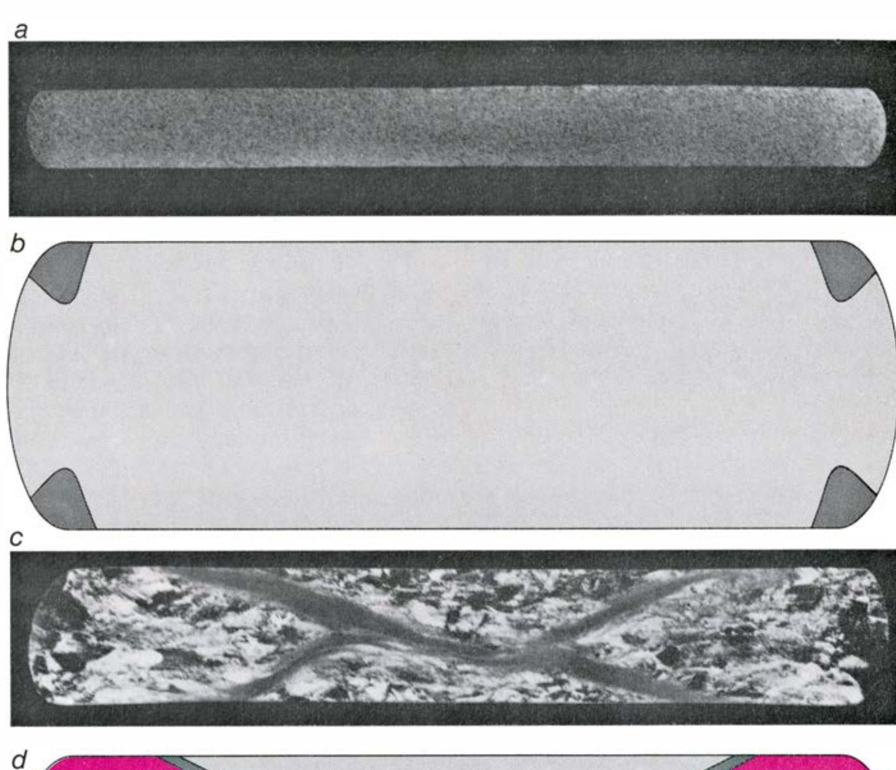
ENTRE LOS DEFECTOS DE LA FORJA EN CALIENTE se incluyen las bandas de deslizamiento que pueden desarrollarse cuando las matrices frías entran en contacto con el metal caliente. Se presentan aquí algunos ejemplos (*a, b*) en secciones transversales de barras forjadas en caliente de una aleación de titanio, precalentadas a una temperatura a la que la resistencia de la deformación es muy sensible a la temperatura. En una prensa mecánica, el hechurado rápido conduce a pequeñas cantidades de calor transferido entre las matrices y el metal,y así se producen bandas de deslizamiento modestas (*a*); en una prensa hidráulica el hechurado lento conduce a una mayor transferencia de calor y, por tanto, a unas bandas de deslizamiento más severas y a grietas por deslizamiento (*b*). Por el contrario, puede conseguirse una deformación relativamente uniforme (*c*) si la pieza de trabajo se precalienta a una temperatura a la que la resistencia a la deformación apenas se altere y si la forja se realiza rápidamente, como ocurre en una prensa mecánica. El diagrama (*d*) representa una estructura no uniforme que contiene bandas y grietas de deslizamiento (*en tono gris*) y zonas de metal muerto más frías (*en color negro*).

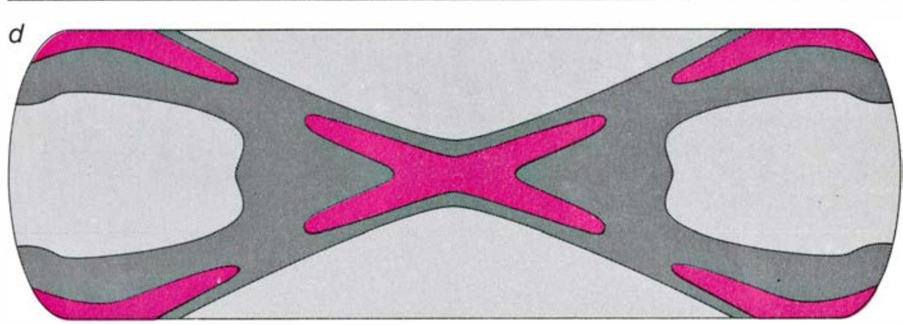
ter Lee, de la Timken Company, y Howard A. Kuhn, de la Universidad de Pittsburgh, han desarrollado unas soluciones para evitar la rotura, que son de gran utilidad en el diseño de piezas forjadas en frío.

Otros tipos de rotura y defectos aparecen en la forja en frío, pero son tan poco frecuentes o de efecto despreciable sobre el servicio del componente que, generalmente, se desprecian. Me refiero, entre ellos, a la porosidad, por culpa de la cual se producen huecos en razón de los esfuerzos de tracción secundarios generados internamente en la pieza de trabajo. La magnitud de los huecos oscila desde un tamaño microscópico hasta superar varias veces el del grano. Los huecos pequeños no tienen por qué afectar a las propiedades de servicio de la pieza; los grandes, llamados aperturas centrales, generalmente causan la rotura de la pieza de trabajo durante la operación de forja, de modo que la pieza nunca entra en servicio.

Otro defecto que aparece ocasionalmente en la forja en frío es la presencia de bandas de deslizamiento. Se trata de regiones de cizalladura localizada que se desarrollan entre áreas que difieren mucho en su grado de deformación dentro de la pieza. Causa principal de la deformación interna no uniforme es la fricción, capaz de producir el "gripado" del metal que toca las matrices y causar, además, que la carga necesaria para deformarlo aumente considerablemente, muy por encima de la requerida para el resto de la pieza de trabajo. Un área gripada se conoce como una zona de metal muerto. En la forja en frío, las bandas de deslizamiento separan la zona de metal muerto del resto de la pieza, son normalmente bastante difusas debido a la influencia de endurecimiento por trabajo y únicamente en casos muy raros llegan a propagarse lo bastante para provocar una rotura seria de la pieza. Basta mejorar la lubricación o incluir leves modificaciones en el diseño de la matriz o de la preforma para aliviar la situación en la que aparecen bandas de deslizamiento.

En la forja a temperatura media y en la forja en caliente, el desarrollo de huecos internos o bandas de deslizamiento sí constituye un problema mucho más grave. El factor principal que limita la deformabilidad en la forja a temperatura intermedia, al que hemos aludido antes, son los huecos o cavidades que surgen en los límites de grano. En la aparición de tales cavidades interviene el esfuerzo de tracción secundario. Las cavidades se inician por la acción de dislocaciones que se desplazan





FORJA ISOTERMICA EN CALIENTE, técnica desarrollada recientemente. Las matrices se calientan a una temperatura aproximadamente igual a la temperatura de la pieza de trabajo, al objeto de reducir al mínimo posible los defectos. La técnica da lugar a una forja uniforme en una barra de titanio (*a*); una simulación por ordenador del proceso de deformación (*b*) predice que las velocidades de deformación únicamente serán algo mayores en las esquinas que en el resto de la pieza. Esta técnica de forja no siempre produce buenos resultados, como se ve en otra barra de titanio (*c*), donde aparecen bandas de deslizamiento. La causa responsable es el ablandamiento por fluencia. La simulación por ordenador, utilizada para predecir la localización de la fluencia (*d*), muestra grandes velocidades de deformación y grandes gradientes, tal como se observó en la forja. En las simulaciones, el intervalo de velocidades de deformación oscila desde valores bajos (*gris*) hasta valores altos (*color*), pasando por valores intermedios (*negro*).

a través de las redes del cristal y chocan en los límites de grano. Si existen partículas microscópicas en los límites, las dislocaciones pueden inducir una acción de deslizamiento entre ellas, abriendo pequeños espacios entre las partículas y los límites de grano. Los esfuerzos de tracción secundarios originan el crecimiento de las cavidades. Si la deformación impuesta es suficientemente grande, las cavidades chocan entre ellas y producen la rotura.

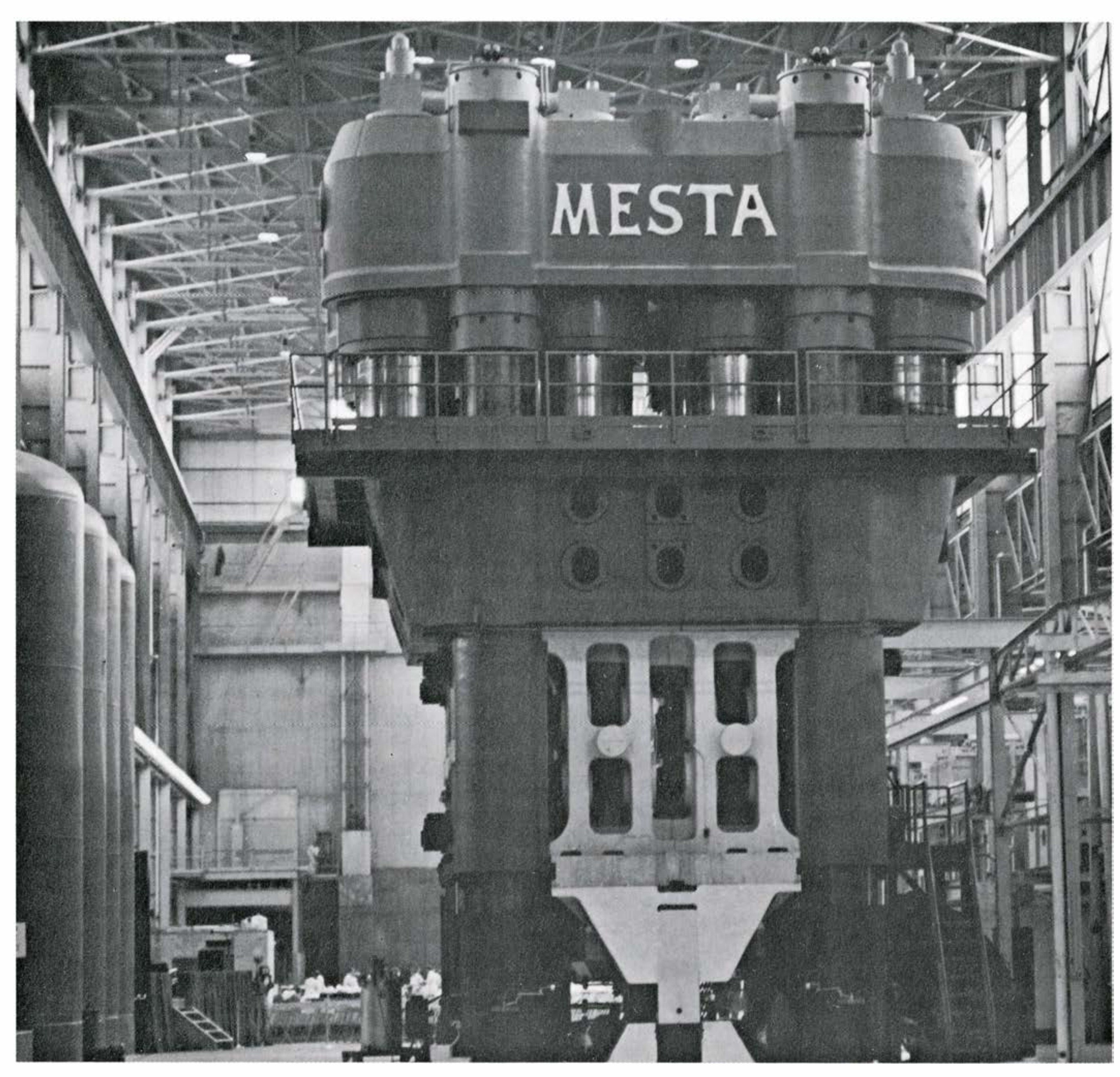
Un segundo tipo de cavidad, que se observa tanto en el trabajo a temperaturas medias como en el trabajo a temperaturas bajas, surge en las uniones triples: lugares donde se reúnen tres granos. A esas temperaturas, la cizalladura a lo largo de los límites de grano se activa térmicamente. Por ser moderadamente altos los niveles de tensión, la cizalladura puede conducir a una concentración de tensiones y cavidades en las uniones. La deformación de la red resulta insuficiente para rebajar la concentración de tensión, y, así, se inician grietas en las uniones que atraviesan longitudinalmente una serie de límites de grano, pudiendo terminar en una grave rotura de la pieza.

Basta, a veces, elevar la temperatura de forja a un régimen de trabajo en caliente para solucionar este tipo de problemas de deformabilidad que hemos descrito. Debe operarse, pues, con sumo cuidado porque el esfuerzo de trabajo en régimen de forja en caliente es bastante sensible a la temperatura: conforme aumenta la temperatura de la pieza de trabajo baja notablemente la resistencia a la fluencia. Este efecto resulta de particular importancia en la forja de ciertos metales, como las aleaciones de titanio utilizadas en los moto-

res a reacción y en la estructura de las alas de los aviones. Cuando la tensión se torna extremadamente sensible a la temperatura, pueden formarse bandas de deslizamiento durante la operación de forja. Nuestro colega Albert Hoffmanner, del Battelle Memorial Institute Columbus, Alan H. Rosenstein, de la Air Force Office of Scientific Research, y los que firman este artículo se cuentan entre los primeros en reconocer el origen de estos defectos y en establecer métodos para evitarlos.

La fuente primaria de bandas de deslizamiento en forja en caliente está en la transferencia de calor entre las matrices y la pieza de trabajo. Esto da lugar a que la pieza se enfríe rápidamente en contacto con las paredes de la matriz, normalmente mucho más frías, originando un aumento del esfuerzo de trabajo y de la carga necesaria para deformarla en comparación con las otras partes de la pieza de trabajo. Se forman zonas de metal muerto y bandas de deslizamiento.

Para impedir la aparición de bandas de deslizamiento en la forja en caliente convencional, se suele subir la temperatura de precalentamiento de la pieza de trabajo hasta un nivel donde el esfuerzo no se halle sujeto a la temperatura. Otro procedimiento consiste en reducir la transferencia de calor, introduciendo lubricantes, que actúan de aisladores térmicos entre las matrices y la pieza de trabajo, o reduciendo el tiempo de contacto forjando con mangos más rápidos. Se pueden calentar también las matrices a la misma, o casi la misma, temperatura de la pieza de trabajo. Se está prestando especial atención a esta última técnica, conocida como forja en caliente isotérmica. En ausencia de cantidades significativas de calor transferido, las piezas se pueden forjar en caliente mediante prensas hidráulicas lentas. Debido a la baja velocidad del mango, las cargas de trabajo son más pequeñas que las aplicadas en la forja en caliente convencional mediante prensas mecánicas o martillos. Además, el intervalo de temperaturas a las que se puede trabajar el metal es considerablemente mayor que en la forja convencional en caliente. Esta



PRENSA HIDRAULICA de la planta de Alcoa en Cleveland (Estados Unidos). Tiene una capacidad de 445 meganewton (50.000 toneladas). La prensa está impulsada mediante bombas hidráulicas y la presión de aceite se mantiene constante gracias a los tanques acumuladores (*a la izquierda de la foto*).

técnica debe aplicarse con la debida precaución a aquellos metales que presentan ablandamiento por deformación durante su trabajo en caliente. En tales metales, propicios a la deformación, se producen forjados defectuosos que contienen bandas y grietas de deslizamiento.

Ante la demanda de forja que se hace en la actualidad, el diseño de preformas y matrices resulta con frecuencia de excesiva complejidad para permitir la aplicación de métodos simples y generales que puedan predecir las cargas de trabajo, el movimiento del metal y el potencial de defectos. Hasta hace poco, la solución a estos problemas de diseño dependía de métodos de tanteo en manos de gente experta. Tales métodos son caros y consumen tiempo, y, a menudo, requieren operaciones experimentales muy elaboradas y mediciones complejas. Por cuya razón se ha puesto un esfuerzo cada vez mayor en la investigación operativa de técnicas matemáticas, mediante las cuales poder simular operaciones de forja con auxilio de los ordenadores. Este es uno de los objetivos principales que se han señalado nuestros colegas del Cooperative Center for Metal Processing, en los laboratorios del Battelle Columbus. El trabajo está dirigido por Taylan Altan, quien fue pionero en la aplicación de los modelos con ordenador a los problemas de forja.

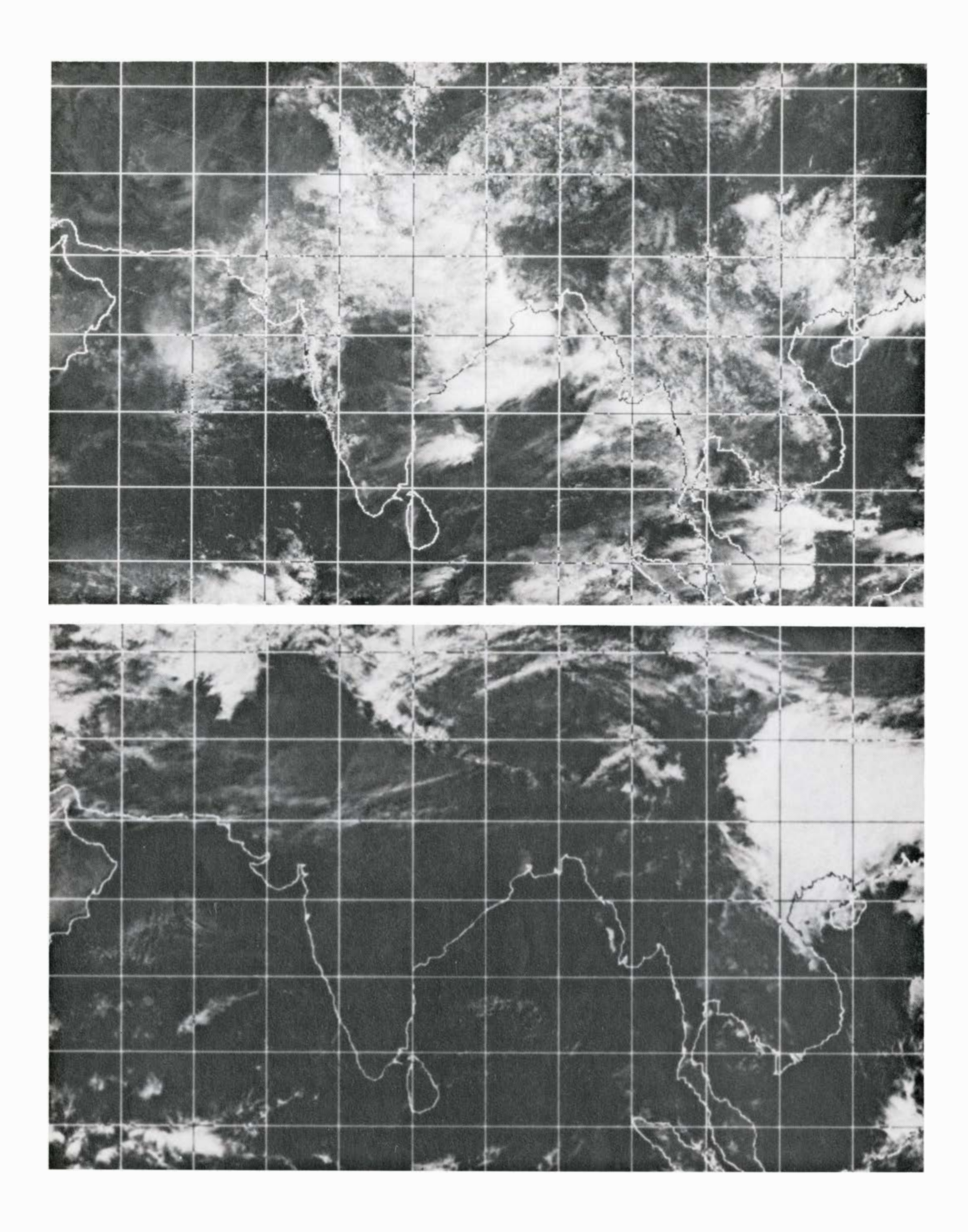
En el pasado, los modelos matemáticos de los procesos de forja se realizaban por varios métodos de aproximación que eran demasiado específicos y fallaban a menudo al dar estimaciones precisas de las fuerzas y energía requeridas. En los últimos quince años se ha aplicado un procedimiento mejor: el método de los elementos finitos, que se ha aplicado con éxito a una amplia gama de problemas de forja. En este método, una sección transversal del material a deformar se divide en un enrejado bidimensional de elementos discretos; la deformación en puntos seleccionados se determina aplicando algún principio variacional. A través del análisis de un agregado de elementos, uno por vez, se puede hallar el tipo de deformación de una pieza de forma complicada. La exactitud de los resultados aumenta con el número de elementos.

Hasta hace poco, la aplicación del método de los elementos finitos a problemas prácticos de forja tenía una barrera infranqueable: para progresar necesitaba un ordenador grande y un operario experimentado. Actualmente, en Battelle, Soo-Ik Oh ha desarrollado un programa para los usuarios. Este puede resolver problemas de forja para componentes de formas complicadas y no requiere un tiempo excesivo. Además, su simulación de la forja proporciona mapas detallados del esfuerzo, de la deformación y de campos de temperatura en la pieza de trabajo y en las matrices. Hemos empleado el programa para predecir las bandas de deslizamiento en forja en caliente isotérmica. Harold Gegel, del Air Force Wright Aeronautical Laboratory, aplicó el programa a la predicción de los esfuerzos de tracción secundarios y la formación de cavidades en la forja en caliente de discos compresores de titanio para motores de reacción.

En el diseño de matrices, también se recurre a los programas de ordenador. Los diseños tienen en cuenta las correcciones necesarias a la contracción térmica del metal en forja en caliente, la deformación elástica de las herramientas y otros factores. Los modelos con ordenador pueden predecir los esfuerzos en áreas críticas de las matrices y, por consiguiente, advertir al diseñador los lugares donde puede fracturarse una matriz. Los modelos también permiten predecir dónde puede no ser uniforme la deformación o dónde pueden desarrollarse cavidades; y, en razón de ello, cambiar el diseño de la preforma.

Con la llegada de los gráficos de interacción (la técnica que permite hacer aparecer y modificar los dibujos de los diseños en una pantalla de un terminal de ordenador), el diseñador puede ver rápidamente los resultados de sus ideas. Y es de ayer mismo otra innovación: el trabajo de los ingenieros de diseño y de los fabricantes de matrices se ha hecho algo más fácil merced a la introducción de los ordenadores-diseñadores. Un sistema tal comprende un ordenador independiente y un equipo avanzado que efectúa el diseño geométrico y lo muestra en la pantalla. Cuando el diseñador ha realizado un plano para las matrices sobre la pantalla, el sistema produce un conjunto completo de dibujos de ingeniería y una cinta magnética, que puede enviarse a la fábrica para mecanizar las matrices.

Los programas con ordenador, que diseñan los equipos de forja y simulan las operaciones de la misma, suministran a los ingenieros un instrumento de gran utilidad y eficacia. Al aplicarse los programas de ordenador, se eliminarán los viejos métodos de tanteo y harán posible mejores diseños de preformas y matrices. Los usuarios de modelos con ordenador podrán predecir con precisa exactitud la fluencia del metal y los potenciales defectos de forja.



CIRCULACION DEL MONZON DE VERANO sobre la India y Sureste de Asia. La corriente transporta aire húmedo del océano ecuatorial al continente (*arriba*). La circulación está producida por diferencias de presión atmosférica entre el aire caliente situado sobre tierra firme y el aire fresco sobre el océano. La intensidad de la circulación se incrementa cuando el vapor de agua transportado por el aire se condensa y libera energía. En la estación seca o invierno, la circulación se invierte y la cobertura nubosa sobre la tierra se hace mínima (*abajo*). Los monzones no son exclusivos del océano Indico; se desarrollan allí dondequiera que haya una variación estacional del viento motivada por calentamiento diferencial de la atmósfera. Las dos fotografías son imágenes compuestas, generadas por ordenador a partir de datos recogidos por el satélite meteorológico *NOAA-5* el 3 de agosto y el 1 de diciembre de 1977.

Los monzones

Proyectando sobre la tierra la energía solar que alcanza el océano, estos vientos estacionales proveen de agua a la mitad de la población de nuestro planeta. Las simulaciones mediante ordenador predecirán pronto sus fases secas y lluviosas

Peter J. Webster

Si la Tierra fuese un planeta más sencillo, cuyo Hemisferio Norte estuviera recubierto en su mayor parte de un continente único, y el resto fuera un vasto océano, las características del tiempo en la vecindad de las costas del continente no diferirían mucho, quizá, de las que regulan la vida en tres de los cinco continentes reales. Habría dos estaciones principales en la llanura costera, la lluviosa y la seca. Dentro de la estación lluviosa, alternarían, cada una o dos semanas, períodos de chubascos intensísimos con otros de tiempo soleado. Los habitantes de las regiones costeras y del interior meridional del hipotético continente llegarían a acostumbrarse a ciclos de cambio estacional regulares y definidos.

En un hecho importante, al menos, la vida en ese planeta hipotético diferiría completamente de la vida en la Tierra real: en el planeta hipotético podrían predecirse con exactitud los principales cambios meteorológicos. La verdad es que ese planeta supuesto constituye un modelo simplificado de la Tierra, que se ha desarrollado, con ayuda de un ordenador, para simular las características globales del tiempo. En el planeta modelo puede predecirse el comienzo de la estación lluviosa, la alternancia de tiempo "durmiente" (seco) y "activo" (lluvioso) dentro de la estación de lluvias y la fecha aproximada de la terminación de éstas al comienzo de la estación seca. Un agricultor del planeta modelo que tuviera acceso a tal información podría planear la época de siembra, seleccionar el tipo de cultivo acorde con la precipitación adecuada y optimizar la probabilidad de una buena cosecha. Semejante predicción incidiría profundamente en la vida diaria de los dos mil millones de personas de la Tierra real que dependen del agua de las lluvias estacionales para consumo humano y usos agrícolas.

El comportamiento del tiempo a gran escala, que se simula en el planeta modelo, define al monzón. Término que se ha venido aplicando a las variaciones estacionales registradas en las costas del océano Indico y, de forma particular, a cierto sistema de vientos del mar Arábigo que soplan del Suroeste durante una mitad del año y, del Nordeste, durante la otra mitad. Se asocia su etimología a la palabra árabe *mausim*, que significa estación. A medida que se han ido desvelando los mecanismos que originan los monzones, el término ha venido a designar cualquier ciclo climático anual con variaciones estacionales del viento que, en general, producen veranos húmedos e inviernos secos. Sin embargo, los mayores monzones, y los más fuertes, se presentan en las regiones de la Tierra donde nació su nombre: en los continentes de Asia, Australia y Africa y en los mares y océanos adyacentes.

Aunque la característica definidora del monzón es una variación estacional, se observan fluctuaciones en escalas de tiempo que oscilan desde días hasta decenios. Las variaciones a corto plazo incluyen no sólo las fases activa y durmiente de la estación lluviosa, sino también las propias perturbaciones singulares de la fase activa. Durante una fase activa, el tiempo es variable, con frecuentes temporales que comportan las inundaciones asociadas, a menudo, a los monzones. A lo largo de una fase durmiente, el tiempo, que es seco, caliente y soleado, se distingue por la ausencia de temporales tropicales. A lo largo de períodos mucho mayores hay variaciones en la precipitación anual que pueden conducir a años de sequía o de inundaciones. Los ciclos de duración superior a un año no se conocen lo suficiente para avanzar predicciones fundadas, aunque pueden esperarse años de inundaciones o de sequía unas 30 veces por siglo. Hay, sin embargo, aportaciones recientes en la teoría de los procesos húmedos de la atmósfera que permitirán pronto predecir las fases activa y durmiente.

El conocimiento práctico y la previsibilidad general del fenómeno de los monzones desempeñaron un importante papel social y económico en muchas de las antiguas civilizaciones del Hemisferio Oriental. Mucho antes de la llegada de los europeos, los mercaderes habían abierto rutas comerciales entre Asia y Africa oriental, adaptando su comercio a los ritmos estacionales. En 1498, un piloto árabe mostró al explorador portugués Vasco da Gama la ruta comercial a la India desde la costa oriental de Africa, y los vientos monzones constituyeron la base de un lucrativo comercio e intercambio cultural entre el Este y el Oeste. Los mercaderes y aventureros europeos regresaron a sus países con informaciones fragmentarias acerca de los vientos estivales del Suroeste y los invernales del Nordeste.

Provistos de tales observaciones meteorológicas correspondientes a las bajas latitudes, los estudiosos europeos estuvieron por primera vez en condiciones de considerar la circulación de la atmósfera a escala global. Dos de los más importantes estudios primitivos los realizaron, a finales del siglo XVII y a comienzos del XVIII, Edmund Halley y George Hadley. Halley atribuyó la circulación monzónica, fundamentalmente, al calentamiento y enfriamiento diferencial de las tierras y el océano. Según su razonamiento, el calentamiento diferencial produciría diferencias de presión en la atmósfera, que serían igualadas por los vientos. Hadley hizo notar que la rotación de la Tierra alteraría la dirección de tales vientos, determinando que los que se movían hacia el Ecuador se desviasen a la derecha en el Hemisferio Norte y, a la izquierda, en el Hemisferio Sur. Aunque los trabajos más recientes han refinado el conocimiento de ambos procesos, éstos se consideran todavía como las causas fundamentales de los fenómenos monzónicos.

Hay, sin embargo, un tercer factor que determina muchas de las caracte-

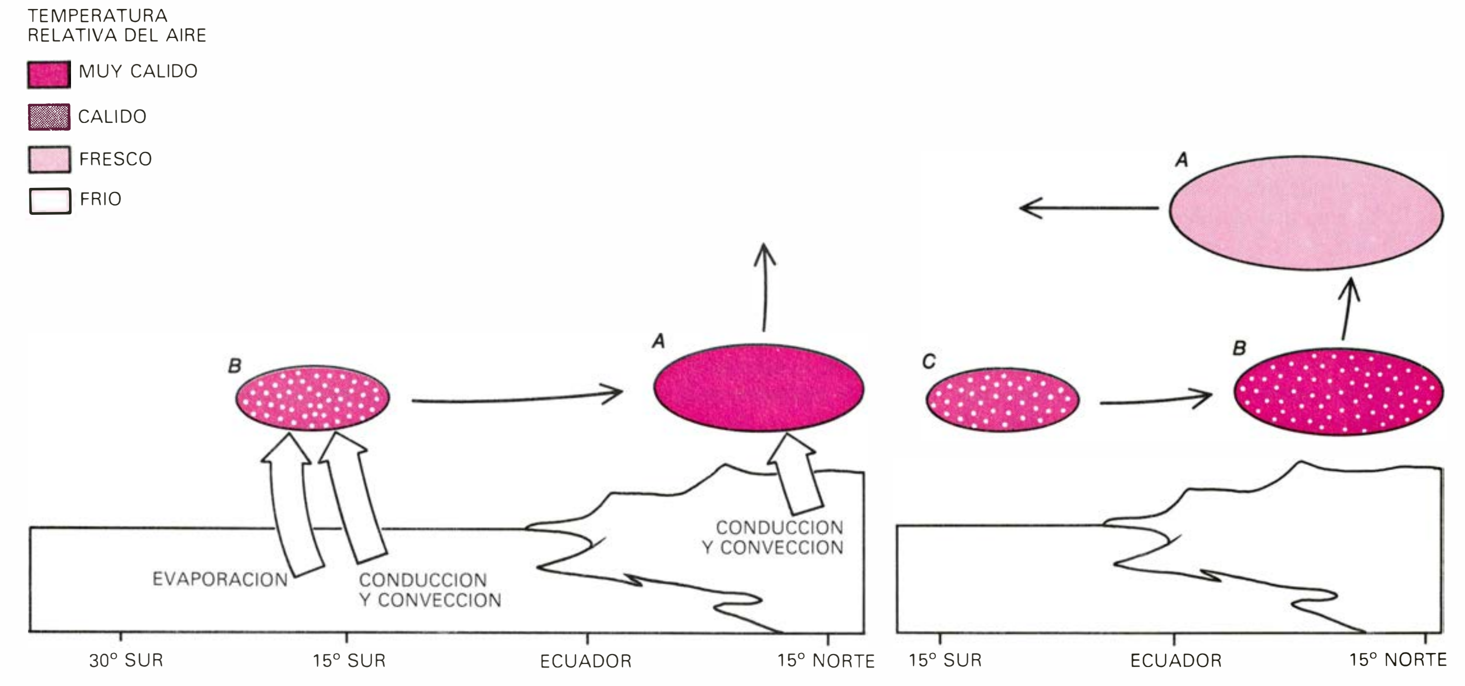
rísticas distintivas de los monzones. El margen de temperaturas y presiones sobre gran parte de la superficie terrestre queda en la vecindad del punto triple del agua. El punto triple de una sustancia es la combinación de temperatura y presión para las cuales pueden coexistir sus fases sólida, líquida y gaseosa. Para el agua, el punto triple se encuentra a la temperatura de 0,01 grados Celsius y a la presión de 6,104 milibares. Las moléculas de agua cuyo ambiente circundante está próximo al punto triple pueden pasar libremente de uno a otro de los tres estados. Por el contrario, las moléculas de dióxido de carbono no pueden pasar al estado líquido a las temperaturas y presiones que ordinariamente se encuentran en la superficie terrestre; la presión debe elevarse considerablemente para que el dióxido de carbono sólido (nieve carbónica o hielo seco) se funda, en lugar de sublimarse a la fase de vapor.

La capacidad del agua para evaporarse y condensarse fácilmente en la atmósfera incide de un modo notable sobre la circulación monzónica. Precisamente en la comprensión de los efectos de los procesos húmedos es donde la meteorología moderna ha aportado su contribución teórica más importante. Cuando el agua pasa de sólido a líquido, debe suministrarse energía para romper la molécula cristalina del hielo, de modo que las moléculas se puedan mover libremente en el estado líquido. Análogamente, se requiere energía para transformar la fase líquida a vapor. La energía consumida para evaporar agua se almacena en forma de energía cinética de las moléculas del vapor de agua; cuando las moléculas se condensan de nuevo, se libera esa energía. Así, durante un cambio de fase, se suministra a una sustancia energía en forma de calor, o se extrae de ella, sin cambiar su temperatura. Puede apreciarse el efecto observando que, aunque el hielo de un vaso de agua congelada se va fundiendo en una habitación caliente, la temperatura del agua no cambia mientras quede algo de hielo.

La importancia de los procesos húmedos obedece a que el agua que en su momento se evaporó de los océanos almacena una sexta parte, aproximadamente, de la energía solar que llega a la superficie de la Tierra. Al condensarse de nuevo y caer en forma de lluvia, se libera la energía almacenada en la fase vapor del agua de lluvia. En la circulación monzónica, parte de la enorme cantidad de energía solar recogida en los océanos se puede liberar sobre el continente cuando el agua del aire húmedo oceánico se condensa sobre las masas terrestres. La liberación de esta energía es la responsable de la potencia y duración de la estación monzónica lluviosa y de la variación, observada dentro de ese mismo período, entre las fases activa y durmiente.

Para conocer con mayor precisión cómo actúan los procesos húmedos en la circulación monzónica, habrá que entender antes bien el modo de operación de los mecanismos básicos impulsores de la circulación; de estos últimos se ocuparía, por vez primera, Halley en el año 1686. La atmósfera no se calienta de una manera uniforme, sino diferencial, debido a que las masas terrestres tienden a estar más calientes que los océanos en verano y, más frías, en invierno. La tierra y los océanos responden de un modo distinto a la radiación solar, por dos razones. Entre las propiedades intrínsecas del agua hallamos su alta capacidad para almacenar calor, si la comparamos con la capacidad de otras muchas sustancias. El calor específico de una sustancia mide su capacidad calorífica: se llama calor específico a la cantidad de energía que debe suministrarse para elevar en un grado Celsius la temperatura de un gramo de la sustancia. En el caso del agua, su calor específico es más del doble que el de la tierra seca, aunque el calor específico de la tierra aumenta considerablemente cuando la tierra está mojada por la lluvia. Por tanto, en respuesta a la misma cantidad de radiación solar, la temperatura de una masa dada de tierra seca aumentará más de dos veces lo que aumente la de una misma masa de océano.

La segunda razón de que sea mayor la capacidad calorífica del océano reside en su eficacia para intercambiar energía calorífica con agua situada a mayores profundidades y distribuir así el calor en una gran masa de agua. El viento que agita la superficie del océano crea remolinos turbulentos que



FASES DE LA EVOLUCION de un monzón de verano. Las diferentes fases están controladas por la interacción de los procesos en aire húmedo con la fuerza que impulsa el aire de mayor densidad hacia regiones donde la densidad es menor. Cuando la radiación solar calienta el continente y el océano, el aire situado sobre ellos se calienta también por conducción y se dilata. Al calentarse el continente antes que el océano, el aire, más caliente, situado sobre la tierra se eleva en forma de burbujas turbulentas y dotadas de fuerza ascensional, y es reemplazado por aire oceánico, más denso. Este transporta

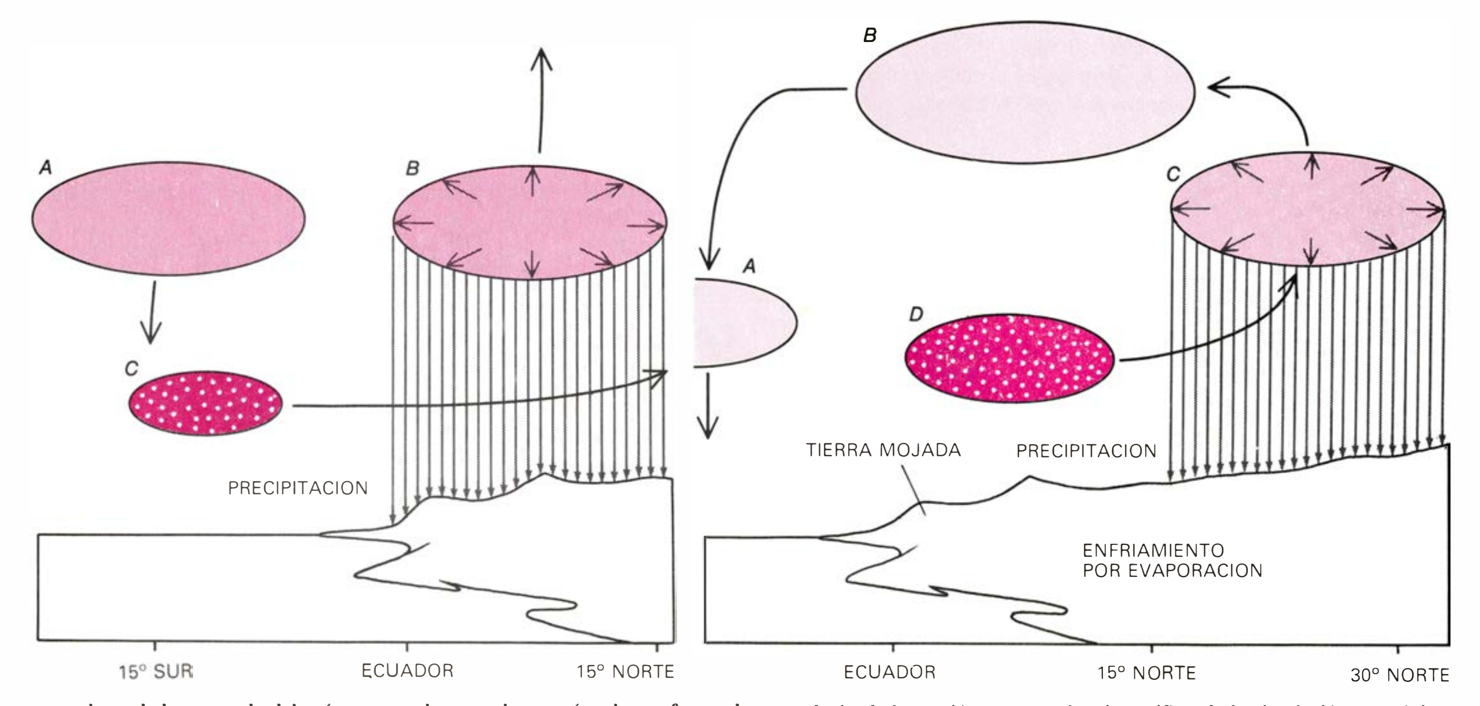
transportan, en verano, agua caliente hasta niveles más bajos; el agua caliente es reemplazada por agua subsuperficial, más fresca, que, a su vez, se calienta. En invierno, el calor acumulado durante el verano se libera mediante el proceso inverso. A medida que la superficie del agua se enfría por la disminución de la radiación solar, el agua superficial se hunde y es reemplazada por agua más caliente, que asciende desde abajo.

En virtud del intercambio y del alto calor específico del agua, la temperatura de la superficie del océano varía menos que la del continente. Los océanos representan enormes volantes de inercia para almacenar energía calorífica, y, como consecuencia de la gran inercia del sistema, el ciclo de temperaturas máximas y mínimas en superficie está retrasado unos dos meses con respecto al correspondiente ciclo de calentamiento solar. En primavera, cuando comienza el ciclo anual del monzón, la energía calorífica que llega a la superficie del océano o del continente pasa a la atmósfera en forma de burbujas turbulentas de aire caliente, dotadas de fuerza ascensional. La velocidad de transporte de calor es proporcional a la diferencia de temperaturas entre la superficie y la atmósfera. A medida que va ascendiendo, la burbuja se mezcla con aire más frío y cede su calor a la columna de aire situada por encima de la superficie recalentada. Esta forma de calentamiento e intercambio de calor se llama calentamiento sensible, debido a que la sustancia calentada ha de estar en contacto con la fuente de calor. El calentamiento sensible causa el calentamiento diferencial inicial de la atmósfera que está sobre la tierra y el océano y genera la energía potencial que impulsa al sistema monzónico.

Los vientos monzones están alimentados por la conversión de parte de la energía potencial del sistema atmosférico en energía cinética. La energía potencial de un sistema sometido a la gravedad es proporcional a la distancia vertical entre su centro de masa y algún nivel de referencia idóneo, la superficie de la Tierra, por ejemplo. La energía potencial puede incrementarse elevando el centro de masa del sistema; esto se consigue aprovechando la energía cinética del sistema, y con ello reduciendo su movimiento, o bien aportando energía procedente de una fuente externa. Por otra parte, si se hace descender el centro de masa, disminuye la energía potencial y, para alimentar el movimiento del fluido, queda disponible una cantidad correspondiente de energía cinética.

Durante el monzón de verano, el calentamiento diferencial aumenta la energía potencial del sistema continente-océano al establecer una diferencia de presión entre partículas de aire sobre uno y otro. El aire que inicialmente está sobre el océano es más frío, y más denso por tanto, que el aire situado sobre el continente. La fuerza generada por el gradiente de presión, que tiende a igualar las diferencias de presión, obliga a que el aire más frío y denso del océano se mueva hacia la tierra, penetrando por debajo del aire caliente situado encima de ésta. Por tanto, el aire caliente se ve forzado a ascender. La combinación del ascenso del aire caliente y del descenso del aire frío hace descender el centro de masa del sistema atmosférico, y así los movimientos del aire liberan energía potencial. Al propio tiempo, la constante aportación de energía solar tiende a incrementar la energía potencial, debido al persistente calentamiento diferencial de la tierra respecto al océano. Los vientos monzones resultan de la tendencia de la atmósfera a hacer mínimo el gradiente de presión entre el continente y el océano. Desde el punto de vista energético, podemos considerarlos, pues, resultado de la conversión de la energía solar en energía potencial y, después, de energía potencial en energía cinética.

La rotación terrestre, a través de la acción de una fuerza no inercial llamada fuerza de Coriolis, desvía la circulación de los vientos monzónicos. Esta desviación distingue los vientos monzones de las brisas marinas diurnas, que también se originan por calentamiento diferencial. Tales brisas surgen y se amortiguan con excesiva celeridad como para verse muy afectadas por la fuerza de Coriolis. La descripción que Hadley dio de la fuerza de Coriolis basta para vientos que se mueven de los polos de la Tierra al Ecuador, aunque el efecto general puede describirse su-



humedad evaporada del océano, que almacena la energía solar en forma de calor latente (*moteado*). Cuando el aire húmedo avanza tierra adentro, se eleva también, y su vapor de agua se condensa, liberando el calor de condensación. El calentamiento adicional obliga al aire a dilatarse y ascender más, reduciendo la presión en su estela e intensificando la circulación monzónica. La lluvia enfría la tierra porque la evaporación del agua absorbe parte de la energía solar incidente. La región de máximo calentamiento en el suelo se mueve tierra adentro y la región de máxima precipitación la acompaña.

mariamente para vientos que se mueven en cualquier dirección. La fuerza de Coriolis desvía los vientos hacia la derecha en el Hemisferio Norte y, hacia la izquierda, en el Hemisferio Sur. La magnitud de la desviación depende de la latitud en que se mueve el viento: la desviación es máxima en los polos y nula en el Ecuador. Es directamente proporcional al seno trigonométrico de la latitud.

El monzón de verano continuaría, en condiciones ideales, en régimen estacionario hasta que se destruyese o perturbase el balance entre la energía potencial generada por la radiación solar y la liberación de energía potencial por parte del sistema atmósferico. Así, en Asia, el calentamiento solar disminuye sustancialmente después del equinoccio de otoño, y la temperatura del océano adyacente comienza a bajar. Simultáneamente, ciertas regiones del Hemisferio Sur, sobre todo en la proximidad del archipiélago indonésico, se convierten en los centros de máximo calentamiento. A medida que se acorta la diferencia de temperatura entre las masas terrestres asiáticas y la de los océanos circundantes, la energía potencial del sistema se va agotando. Se dice que el monzón se retira, y en el Hemisferio Norte empieza la estación seca.

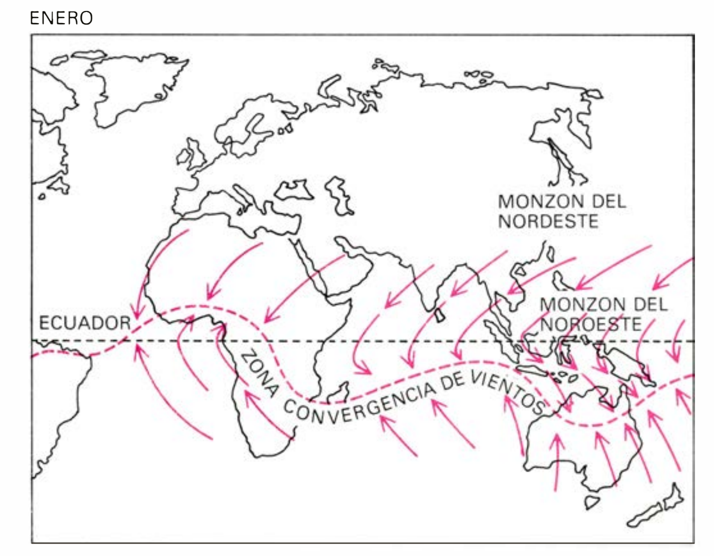
Con el comienzo del invierno, la tierra firme y los océanos del Hemisferio Norte pierden calor por radiación hacia el espacio. Aunque las nubes atenúen las pérdidas radiativas, la cobertura nubosa suele ser menor sobre los continentes que sobre los océanos. Debido a la mayor pérdida de calor por parte de la tierra y a la mayor capacidad calorífica del océano, se restablece una diferencia de temperaturas entre las dos partes del sistema y vuelve a aumentar la energía potencial. El aire frío y a presión alta situado sobre Asia septentrional se mueve hacia el Ecuador para restablecer el equilibrio, y es desviado hacia la derecha por la fuerza de Coriolis. La masa de aire frío, que viene del Nordeste moviéndose sobre la superficie, es compensada por aire caliente del Sur que avanza hacia el Norte, en la alta troposfera. La rotación de la Tierra desvía el aire superior caliente hacia el Este, creando una intensa corriente en chorro sobre Asia y el Japón, cuyas velocidades suelen alcanzar los 100 metros por segundo, o más de 350 kilómetros por hora. La corriente en chorro, que con frecuencia inestabiliza sobre el Pacífico central y septentrional, favorece la formación de las borrascas depresionarias de invierno dominantes en las latitudes medias y altas del Hemisferio Occidental.

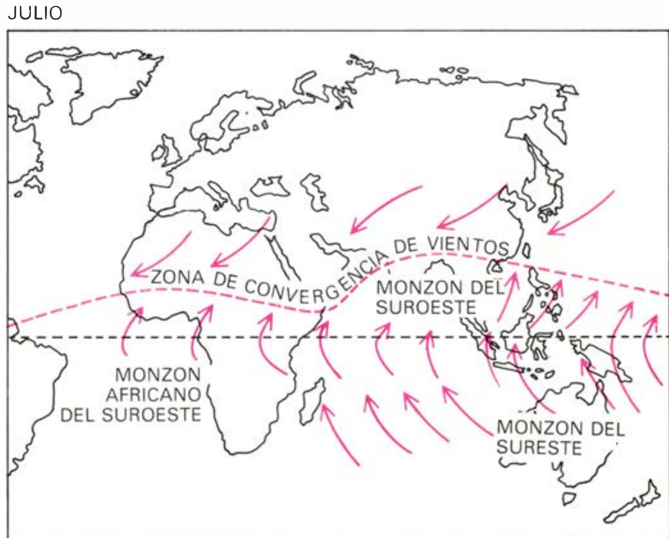
El monzón invernal del Nordeste continúa en régimen estacionario, de manera muy parecida a la del monzón estival del Suroeste, hasta que el calentamiento solar de la primavera disipa la energía potencial que propulsa al monzón. A medida que la temperatura del continente sobrepasa de nuevo la temperatura del océano, comienza a almacenarse energía potencial, y el ciclo empieza una vez más.

¿Qué papel cumplen los procesos de aire húmedo en la circulación anual? Durante el monzón de verano, el vapor de agua evaporada sobre el océano es transportado junto con el aire que se mueve hacia el continente. Si una partícula de aire se mueve verticalmente de suerte que no ceda ni tome energía calorífica, la presión y la temperatura variarán de acuerdo con lo que se llama evolución adiabática. Si la partícula se eleva, avanzará hacia una región donde la presión sea más baja. Se creará entonces un gradiente de presión entre la partícula de aire y su nuevo entorno; a fin de igualar las presiones, la partícula tenderá a dilatarse. Para ello, sin embargo, la partícula habrá de realizar trabajo contra el aire circundante a expensas de la energía cinética de sus moléculas. Reducir la energía cinética de las moléculas implica reducir la temperatura. Si no hay intercambio de calor a través de la superficie que limita la partícula, el proceso se llama enfriamiento adiabático. Inversamente, en el calentamiento adiabático, la temperatura de una partícula de aire descendente aumenta al comprimirla el ambiente, aun cuando no se haya añadido ni sustraído energía calorífica a la partícula.

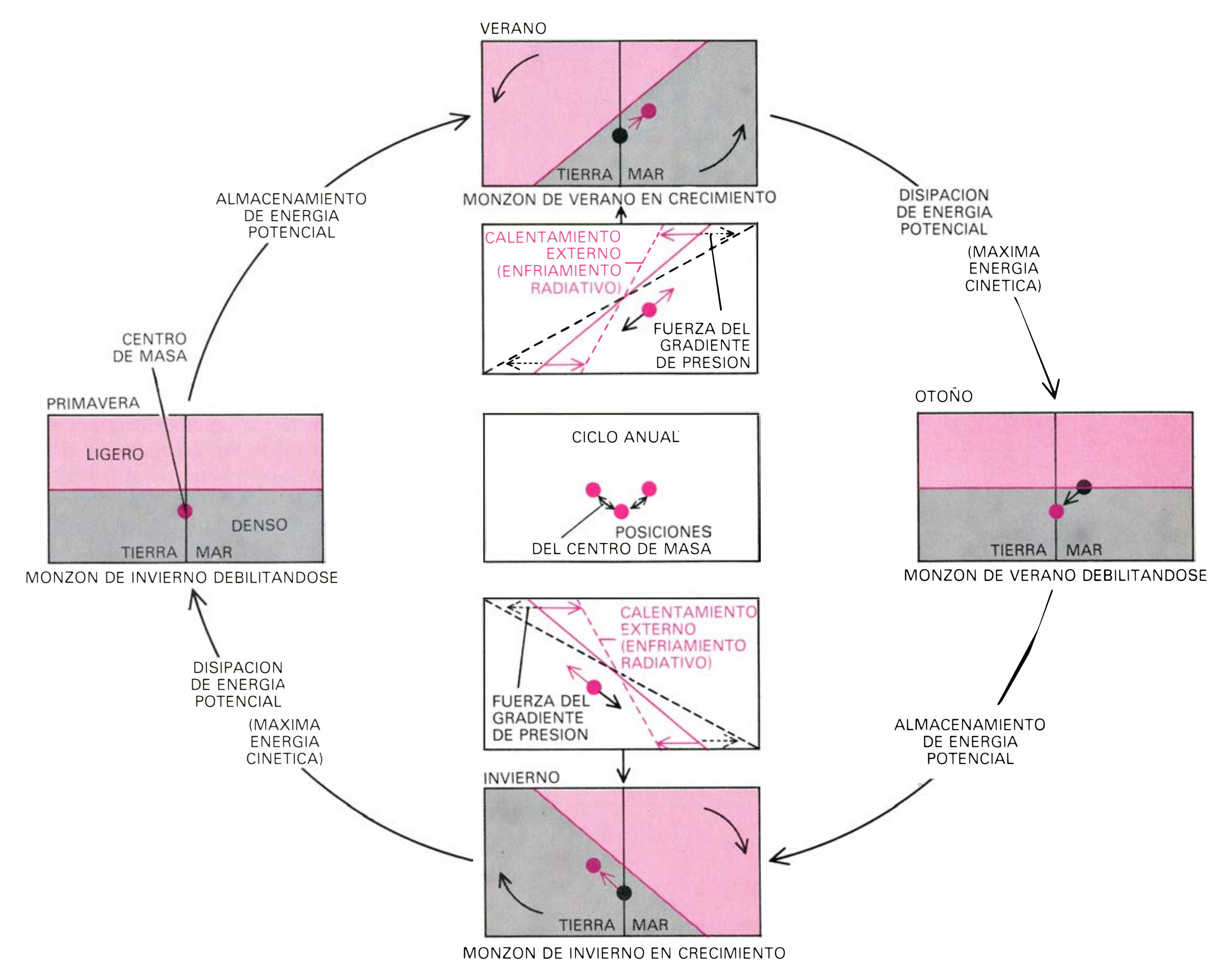
Cuando una partícula de aire que lleva humedad procedente del mar se calienta por conducción y por corrientes convectivas de aire ascendente sobre la tierra, comienza a elevarse hacia alturas donde la presión es menor. Al ascender, se enfría adiabáticamente y el vapor de agua se condensa formando gotas de lluvia. En el curso de la condensación, se libera la energía solar que se utilizó para mantener el agua en la fase de vapor. La energía calorífica liberada, llamada calor latente, es absorbida por las moléculas de aire y, por tanto, produce en la partícula una variación no adiabática de temperatura.

El calor liberado de esta manera añade una considerable fuerza ascensional a la columna de aire ascendente sobre un continente caliente. La partícula de aire asciende a mayor altura, provocan-





CAMBIO ESTACIONAL DEL VIENTO. Característico de la circulación monzónica en todo el mundo, se hace más pronunciado en aquellas regiones que circundan el océano Indico. Obsérvese que, al cruzar el Ecuador, el viento cambia de dirección, efecto que se debe a la rotación terrestre. La zona donde los vientos en superficie convergen se encuentra principalmente en el Hemisferio Sur en enero, pero emigra hacia el Norte, con el Sol, en julio.



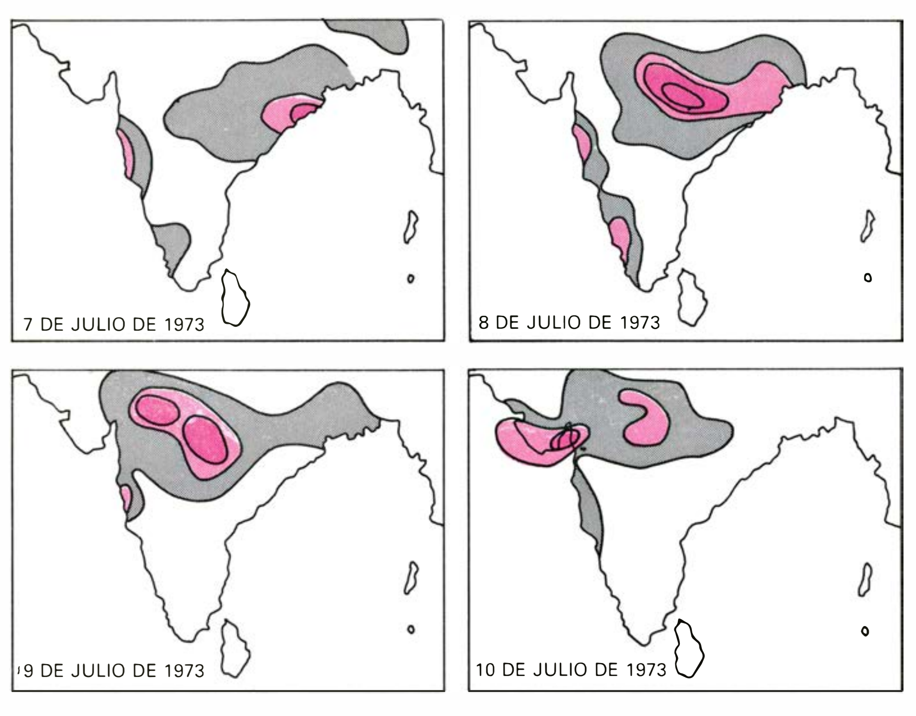
LA DIFERENCIA DE TEMPERATURA entre el aire situado sobre el continente y el que cubre el océano eleva el centro de masa del sistema atmosférico, aumentando así su energía potencial. La diferencia de temperatura se mantiene por calentamiento exterior en verano y por enfriamiento radiativo en invierno. La corriente originada por el gradiente de presión entre las dos masas de aire tiende a disminuir la energía potencial del sistema y a aumentar su energía cinética. La circulación monzónica resulta del equilibrio entre los dos efectos. Cuando el monzón se encuentra en régimen estacionario (verano e invierno), la ganancia de energía potencial por calentamiento externo o por enfriamiento radiativo es exactamente igual a la pérdida causada por la acción de la fuerza resultante del gradiente de presión. En invierno y en primavera el centro de masa desciende y la energía potencial del sistema se disipa.

do una mayor reducción de presión sobre la tierra y dando lugar a una entrada más vigorosa de aire húmedo procedente del océano. De aquí que uno de los efectos de los procesos en aire húmedo sea intensificar la circulación monzónica. Sin procesos húmedos, la circulación continuaría existiendo, pero sería considerablemente más débil.

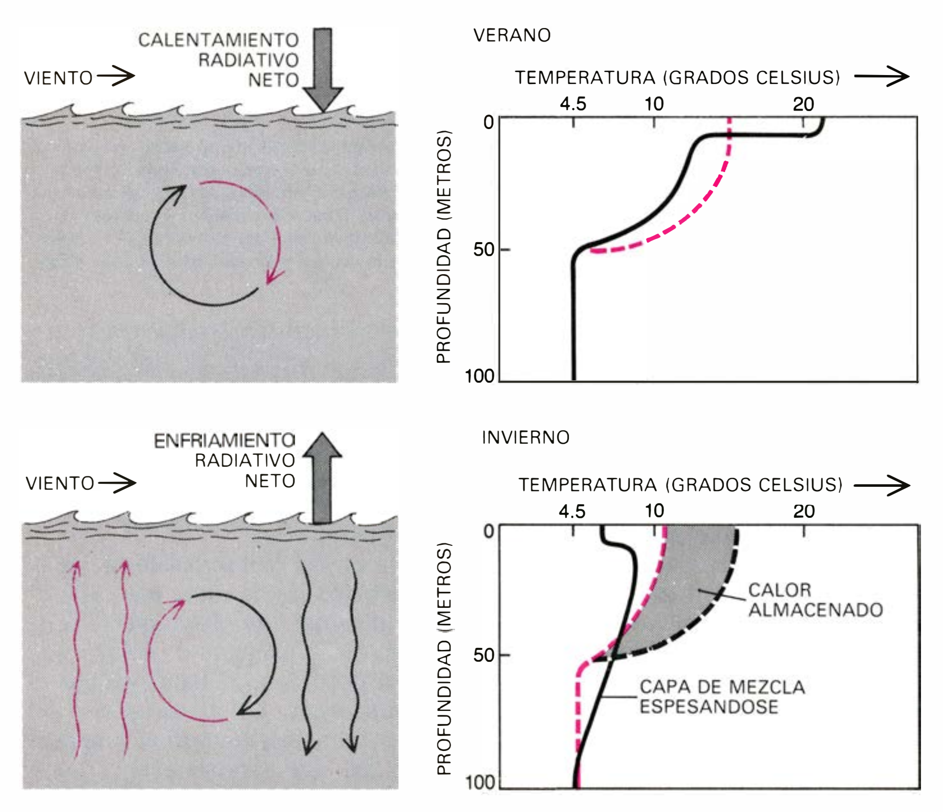
A causa de la columna de aire ascendente, las temperaturas en la alta troposfera sobre Asia meridional son mucho más altas en los subtrópicos que en el Ecuador. El resultado es una fuerza, debida al gradiente de presión, dirigida hacia el Ecuador a altos niveles, opuesta a la fuerza en la proximidad de la superficie. La rotación de la Tierra desvía la corriente en altura hacia el Oeste, dando lugar a una intensa corriente en chorro dirigida de Este a Oeste en la que los vientos pueden alcanzar velocidades de 50 metros por segundo, es decir, 180 kilómetros por hora. Esta corriente en chorro se extiende a través del océano Indico y Africa, donde cruza el Ecuador y se une a los vientos invernales del Oeste del Hemisferio Sur. Los vientos terminan por descender sobre el cinturón subtropical de altas presiones, formado durante el invierno en el Hemisferio Sur. Las características generales de la corriente superior de retorno fueron predichas por Halley basándose solamente en deducciones racionales.

La influencia de las evoluciones en aire húmedo se aprecia, mejor que en cualquier otro aspecto, en sus efectos sobre la fecha de los sucesos que constituyen el monzón estacional. Los procesos en aire húmedo definen, virtualmente, la fecha de comienzo, parecen determinar su máxima intensidad y controlan su retirada. El inicio del período de precipitación llega, bastante bruscamente, algunas semanas después del solsticio de verano. Pero la circulación no alcanza su máxima intensidad hasta unas ocho o 10 semanas después del solsticio.

La razón del retraso reside en que la precipitación media sobre una superficie continental está directamente relacionada con la temperatura del aire que procede del océano. Cuanto más alta es la temperatura de la superficie del océano, tanto más elevada es la temperatura del aire suprayacente y tanto mayor la cantidad de vapor de agua que puede contener. Cuando el aire que sopla del océano lleva más vapor de agua, mayor es la energía que puede liberarse al condensarse sobre la tierra. De ahí que la intensidad de la circulación mon-



FLUCTUACIONES durante la fase activa del monzón de verano. Se producen en virtud del efecto de cizalladura, o variaciones de un lugar a otro en la velocidad horizontal del viento. Las regiones de cizalladura suelen inestabilizarse y se desprenden de la corriente principal de aire, girando como si fueran remolinos turbulentos que se apartaran de la corriente principal de un curso fluvial. Una perturbación de este tipo puede intensificarse por liberación de calor latente en una zona de precipitación y extenderse a centenares de kilómetros. Podemos seguir el curso de la perturbación cartografiando la acumulación diaria de precipitación a lo largo de un intervalo de varios días.



EL CALOR SOLAR que alcanza la superficie del océano durante el verano se distribuye por el interior de una capa de 50 metros de espesor, en virtud de la turbulencia inducida por el viento. En invierno, el ciclo se invierte; el agua fresca superficial, empujada hacia abajo por el viento y por su superior densidad, se reemplaza por agua más caliente y menos densa, acumulada bajo la superficie a lo largo del verano. Las líneas a trazos representan la distribución media de energía calorífica en función de la profundidad. La mezcla es la principal razón de que la temperatura del océano vaya retrasada respecto al ciclo solar.

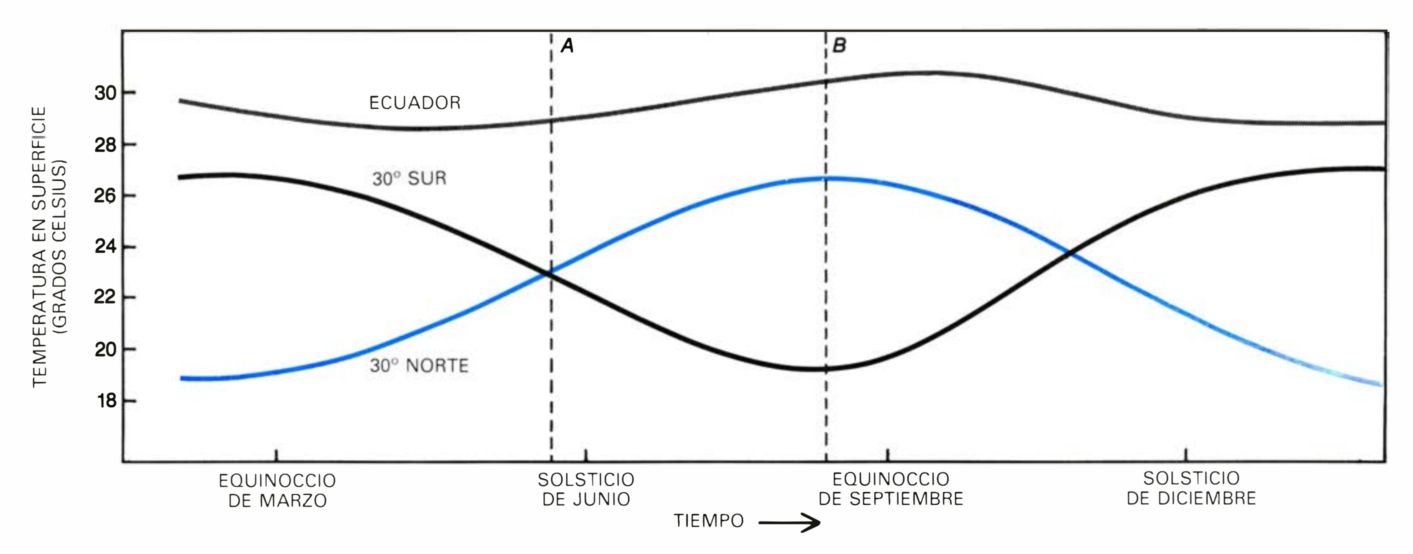
zónica aumente. Los océanos del Hemisferio de verano alcanzan su temperatura máxima unos dos meses después del solsticio de estío; por tanto, la cantidad de vapor de agua transportada por el viento monzónico no llegará a su valor máximo hasta mediados o finales de agosto en el Hemisferio Norte y hasta fines de febrero o comienzos de marzo en el Hemisferio Sur.

La retirada del monzón implica el cese gradual de la precipitación, pasado el equinoccio de otoño. No sólo se reduce, en esa época, el calentamiento diferencial entre los océanos y los continentes, sino que también lo hace la energía inyectada en el sistema por el transporte de vapor de agua. El aire más fresco situado sobre el océano contiene menos vapor, por cuyo motivo decrece gradualmente el calor latente liberado por la precipitación sobre tierra firme.

Con esto se puede ya esbozar la teoría de los procesos en aire húmedo sobre la cual se basa el modelo de ordenador de la circulación monzónica. Recordemos que nuestro planeta hipotético consta de un océano y un continente, que cubre simétricamente la porción superior del Hemisferio Norte. La costa continental coincide con el paralelo 14 grados Norte. En lo demás, el planeta modelo es idéntico a la Tierra: se halla a la misma distancia del Sol, gira alrededor del mismo eje inclinado, con el mismo período y su atmósfera es, física y químicamente, la de la Tierra.

Se supone, en el modelo, que la atmósfera responde a una variación no adiabática de temperatura de una partícula de aire, restituyendo, adiabáticamente, a su valor original la temperatura de la partícula. La hipótesis concuerda estrechamente con el comportamiento de la atmósfera tropical real. Así, cuando a una partícula de aire se aporta calor, de un modo no adiabático, por conducción o como resultado de la condensación de vapor de agua, el aire asciende hasta una región de presión más baja y se enfría adiabáticamente. Por otra parte, si la partícula pierde calor por radiación (proceso no adiabático), el aire desciende y se contrae adiabáticamente hasta que su temperatura retorna al valor inicial.

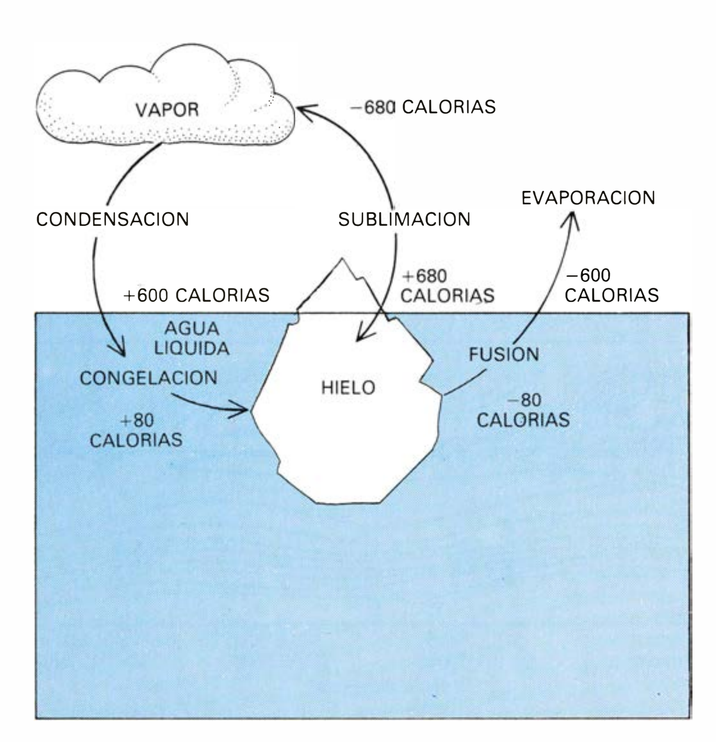
Para simular de forma más verosímil la tendencia de la atmósfera a compensar la aportación no adiabática de calor, hay que tomar en consideración un complicado sistema no lineal. El movimiento vertical del aire sobre un continente recalentado es el resultado de formas interdependientes de calentamiento no adiabático, principalmente



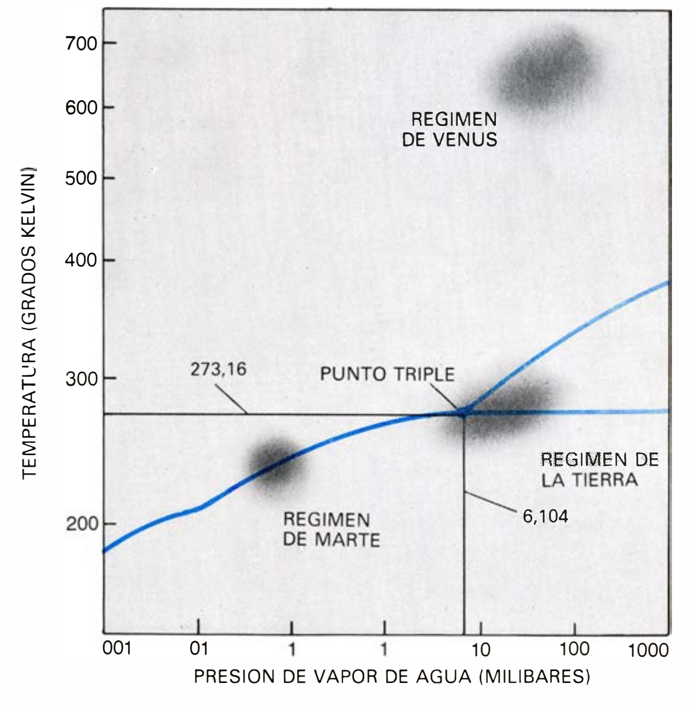
TEMPERATURAS MAXIMAS Y MINIMAS de la superficie del océano. No coinciden con los extremos del ciclo solar, sino que están retrasadas respecto al Sol entre ocho y 10 semanas. Como la energía potencial del sistema monzónico viene determinada, a grandes rasgos, por la diferencia de temperaturas entre los hemisferios Norte y Sur, la energía potencial alcanza un máximo unos dos meses después de los solsticios y un mínimo otros dos después de los equinoccios. La máxima temperatura del océano determina también la máxima velocidad de evaporación; así pues, los procesos en aire húmedo presentan su mayor actividad en la época de máxima energía potencial. Es entonces cuando los monzones son más intensos. La ilustración es de Alan D. Iselin.

calentamiento sensible y latente, además de la respuesta adiabática. Cuando llueve sobre el continente, el suelo se moja, y parte de la radiación solar que se empleaba antes en calentar la superficie va ahora a gastarse en evaporar parte de la humedad del suelo. La tierra se enfría y decrece la cesión de calor sensible al aire situado encima de la tierra mojada.

La reducción de calor sensible no altera de forma importante el mecanismo propulsor de los monzones, porque el calor sensible viene a constituir tan sólo una décima parte del calor latente que se libera en la precipitación. La reducción sí produce, en cambio, el efecto de trasladar la posición del máximo de calor total (calor sensible más calor latente) hacia el interior de la masa continental. La región donde la velocidad vertical de la circulación presenta su máximo sigue el movimiento de la posición de máximo calentamiento, de modo que la condensación de vapor se dirige hacia el interior. La célula monzónica deja en su estela una superficie de tierra saturada de agua y una masa de aire relativamente seca y descendente. El agua acumulada en el suelo se evapora lentamente y la temperatura de la tierra que se está secando comienza a ascender. Al aumentar el calenta-



PARA CAMBIAR LA FASE DEL AGUA de sólido a líquido o de líquido a vapor se requiere una fuente externa de calor. Puesto que la energía se conserva, la energía calorífica queda almacenada en forma de calor latente en la fase que corresponde a una temperatura más alta. En la transición de fase de vapor a líquido, de vapor a sólido o de líquido a sólido, se libera la energía almacenada. La temperatura de una sustancia permanece constante durante un cambio de fase, aun cuando en el proceso hay intercambio de energía.



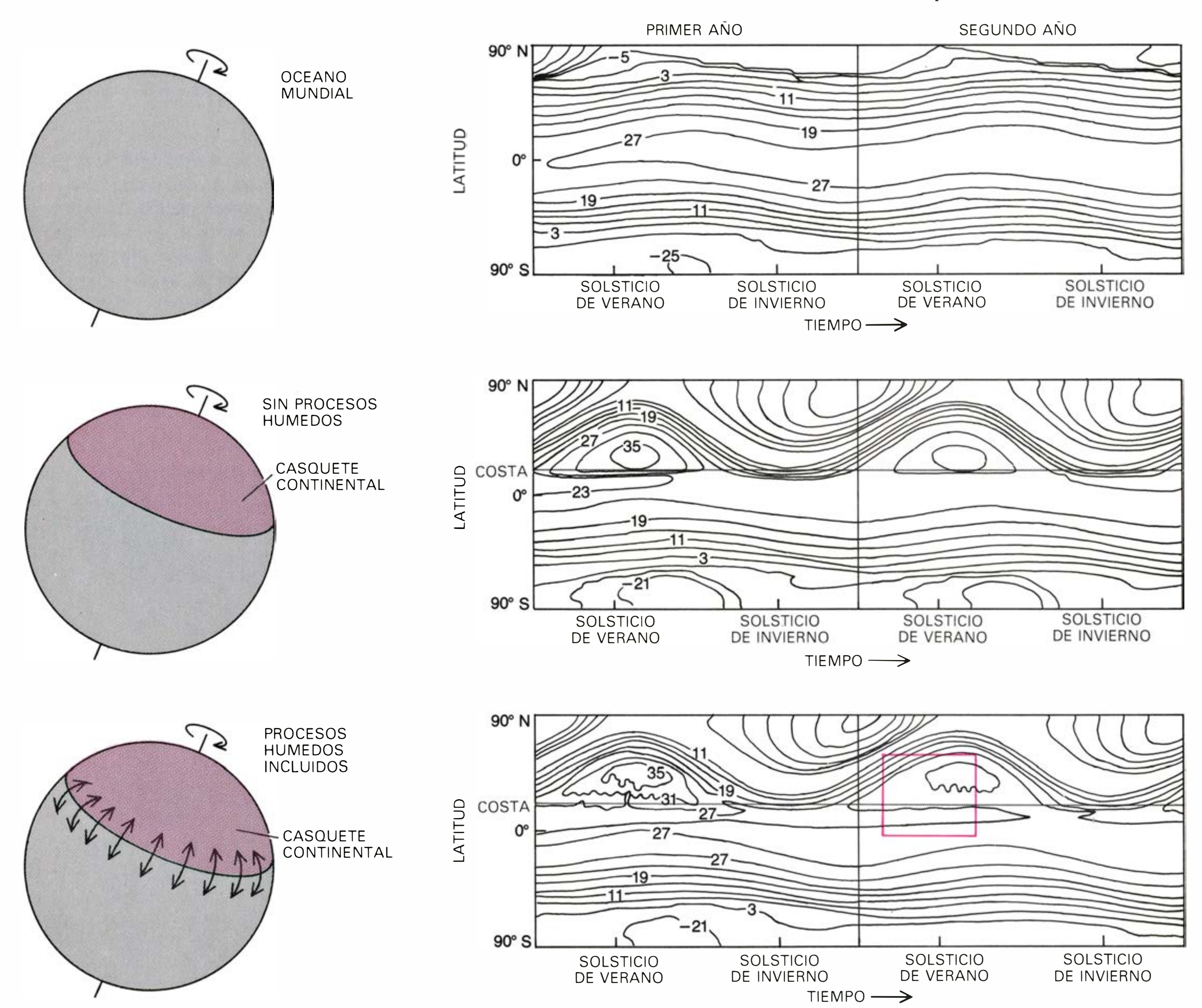
PUNTO TRIPLE DEL AGUA. Se trata de la única combinación de temperatura y presión a la cual pueden coexistir las fases sólida, líquida y vapor del agua. De entre todos los planetas, la Tierra es el único que mantiene condiciones próximas al punto triple. El agua cuyo entorno está próximo al punto triple puede libremente cambiar de una a otra de las tres fases. La energía solar se puede así recoger y almacenar en una fase de mayor energía y puede después transportarse y liberarse sobre regiones relativamente pequeñas.

miento sensible sobre la región continental inmediata a la costa, se forma de nuevo una zona de baja presión y los vientos húmedos procedentes del océano se desvían de su rumbo inicial que era hacia el interior. Una segunda célula monzónica domina la circulación tierra adentro; vuelve a producirse precipitación cerca de la costa. Ciclo que se repite.

En el modelo, la componente vertical del viento sobre una porción del continente puede mostrar la alternancia de las fases activa y durmiente del monzón. Las regiones de ascendencia representan la fase perturbada o activa. La estela, donde domina la subsidencia y no pueden desarrollarse las células tormentosas, constituye la fase durmiente. Si se examina la distribución de temperatura y de la componente vertical de la velocidad del viento a lo largo de varios días, se advertirá que una fase activa del monzón se mueve hacia el interior desde la costa, seguida luego de un período durmiente.

Las observaciones de satélites corroboran, hasta cierto punto, la teoría de las fases activa y durmiente del monzón involucrada en el modelo matemático. Revelan que el lugar geométrico de máxima nubosidad sobre el océano Indico va trasladándose gradualmente hacia el Norte, conforme progresa la célula monzónica. Las perturbaciones del monzón activo y la precipitación intensa están asociadas a la banda de nubes. En la región despejada, detrás de esta banda y sobre la mayor parte de la India, se registra una interrupción de la fase activa que anuncia un período de altas temperaturas, pero sin lluvia. El ciclo tiene un período de 15 a 20 días, en concordancia aceptable con las predicciones del modelo.

Hay tres problemas de predicción distintos, que corresponden a las tres escalas de tiempo del fenómeno monzónico. Para los sucesos más breves, las borrascas de la fase activa, por ejemplo, podemos elaborar analogías matemáticas del sistema físico que produce el mal tiempo. Cuando se le sumi-



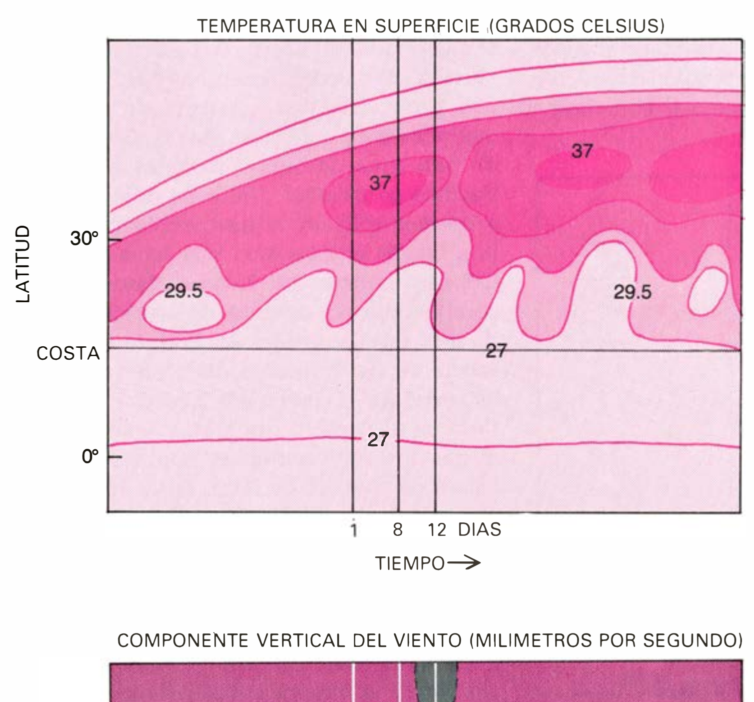
SIMULACIONES DE LOS PROCESOS MONZONICOS realizadas con ordenador usando tres modelos simplificados de la Tierra, que muestran cómo la temperatura en superficie (*isolíneas*) varía con el tiempo y con la latitud. En el modelo de la Tierra cubierta completamente por un océano, la temperatura de la superficie varía de manera muy lenta. Cuando la parte alta del hemisferio del planeta modelo está cubierta por un continente único y en la simulación no se incluyen los procesos en aire húmedo, la evolución de la temperatura del océano se retrasa muchas semanas respecto a la del continente. La máxima temperatura en tierra firme se da justamente al Norte de la costa. Cuando se incluyen los procesos en aire húmedo, la tierra se enfría justamente al Norte de la costa y la temperatura en superficie varía cíclicamente durante el monzón de verano. Se supone que las características del planeta Tierra (tales como la distancia al Sol, la inclinación del eje de la Tierra y la composición de la atmósfera) son las mismas que las que definen al planeta modelo. La simulación puede también mostrar cómo la componente vertical de la velocidad del viento varía con el tiempo y la latitud (no se ha representado aquí).

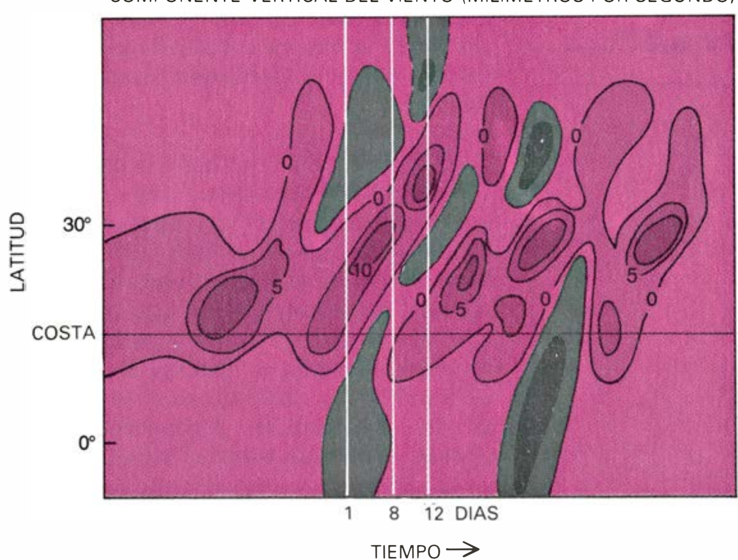
nistran los datos que describen las condiciones actuales del tiempo, el método de simulación digital puede extrapolar los datos hacia el futuro de acuerdo con las reglas de la analogía matemática. La así llamada "predicción numérica del tiempo" constituye una técnica bien establecida que puede alcanzar una razonable exactitud en pronósticos adelantados para varios días.

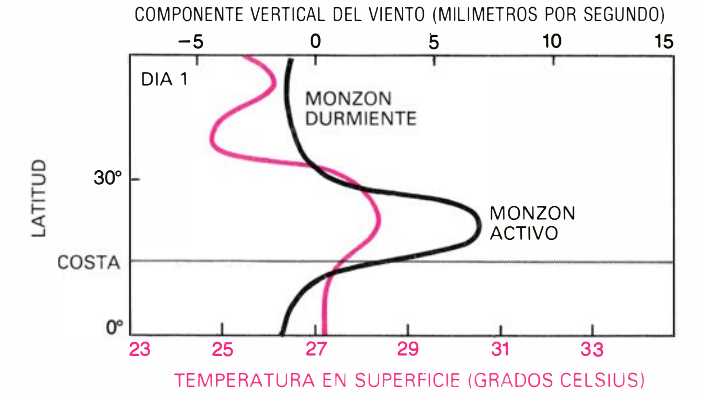
Pero el método presenta un inconveniente muy grave: los datos iniciales deben constituir una descripción completa del estado de la atmósfera en una región dada, incluidas las variables que caracterizan los procesos en aire húmedo. La recogida de tales datos está probablemente fuera del alcance de las redes de observación disponibles. La exploración remota de la atmósfera mediante satélites y boyas flotantes, que operen en colaboración con la red de estaciones de observación en superficie, podrá suministrar en el futuro los datos adecuados para la predicción.

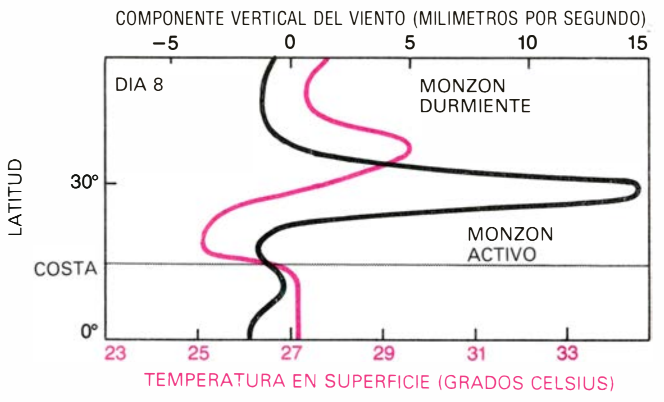
A la hora de pronosticar tendencias estacionales o anuales, apenas se emplean las analogías matemáticas. También aquí los datos escasean, y lo que es más importante, los cálculos resultan demasiado complicados para facilitar predicciones numéricas exactas del clima. No obstante, del estudio de largas series de datos se deduce que ciertas fluctuaciones climáticas están relacionadas con ciertos sucesos precursores. Por ejemplo, el clima invernal de América septentrional parece condicionado por anomalías en la distribución de temperaturas de superficie en el Pacífico Norte. Dada la importancia de los procesos en aire húmedo en relación con los fenómenos monzónicos, quizá valga la pena buscar correlaciones semejantes, por ejemplo, entre fenómenos monzónicos extremos y temperaturas anormales de la superficie oceánica.

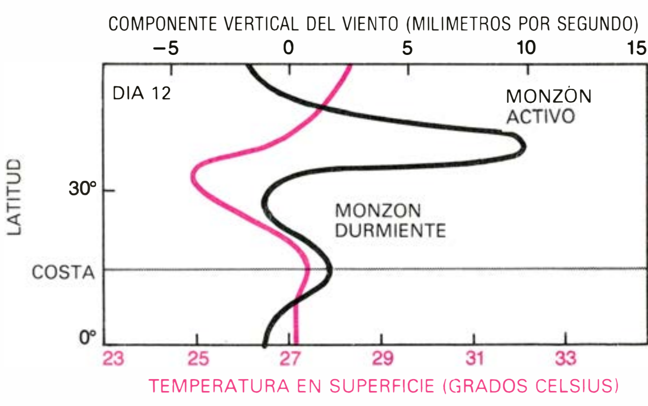
En la predicción de los sucesos a escala de tiempo intermedia, es decir, las fases activa y durmiente del monzón, es donde la predicción podría tener su mayor impacto económico y social. Predecir la transición de una fase a otra requiere un pronóstico para un período









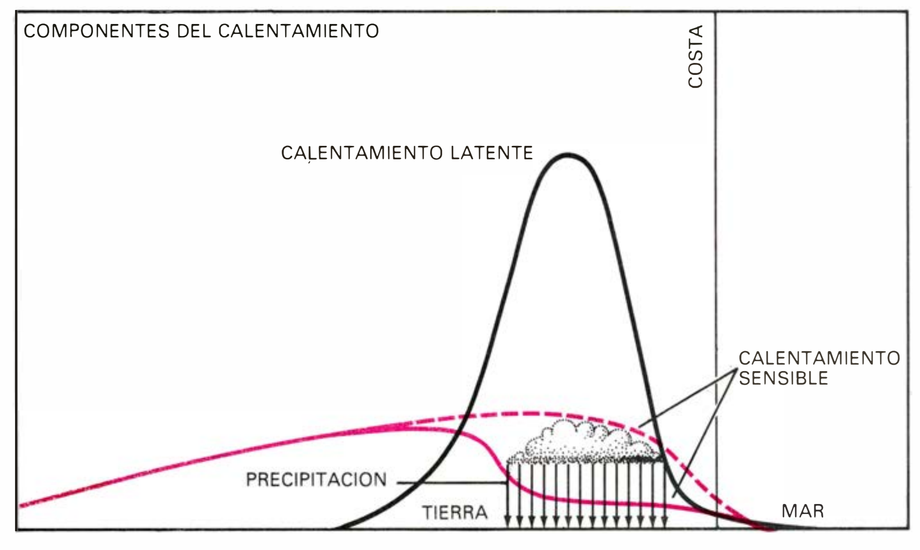


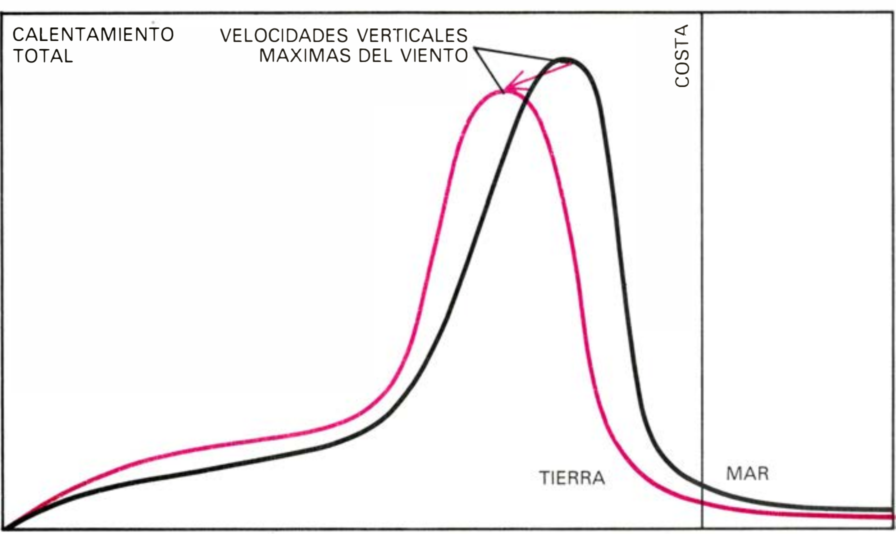
FASES ACTIVA Y DURMIENTE del monzón de verano en la simulación matemática. Se producen cuando los procesos de aire húmedo se incluyen en el modelo hipotético con un casquete continental. La distribución de la temperatura en superficie con el tiempo y la latitud es una ampliación del área enmarcada en color en el diagrama inferior de la página opuesta. También se representa la distribución de la velocidad vertical del viento. Las corrientes de aire ascendentes (*zona coloreada*) son características del monzón activo, mientras que el aire descendente (*región gris*) inhibe el desarrollo vertical de nubes y corresponde al monzón durmiente. En el diagrama de la derecha se representa la distribución de temperatura y velocidad vertical del viento en un corte vertical para tres días del ciclo. Para hacer patente el paso de las fases activa y durmiente, el gráfico de cambios de temperatura muestra la diferencia entre la temperatura real y la temperatura media para esa latitud a lo largo de un período de 30 días. La temperatura desciende en la estela del monzón activo en los días ocho y 12 y sube cerca de la costa en el día 12, a medida que se seca el suelo. La fase activa del ciclo comienza de nuevo.

de semanas, demasiado largo para técnicas numéricas exhaustivas. Pero, si son correctas las teorías de la alternancia de fases activa y durmiente, podría superarse el escollo. Se podría tal vez desarrollar una analogía matemática más sencilla y eficaz, que ignorase los sucesos a escala de tiempo más corta y se concentrase en los elementos del sistema de variación más lenta.

Dentro del contexto de la teoría general de la circulación monzónica, hay variaciones locales y continentales que deben tenerse en cuenta. El Himalaya y la altiplanicie tibetana, por ejemplo, parecen acelerar el comienzo del monzón asiático y aumentar su intensidad final. Datos obtenidos mediante satélites indican que las partes central y suroriental del Tíbet permanecen libres de nieves durante la mayor parte del año. Por tanto, la meseta se calentará rápidamente durante la primavera del Hemisferio Norte. A pesar de haber sido objeto de múltiples estudios teóricos, no acaba de saberse con exactitud cómo influye la meseta en la atmósfera.

¿Qué ocurre con los monzones de verano de Australia y Africa? Sencillamente, que tienden a ser bastante más débiles que su contrapartida asiática. Sobre Australia septentrional, la precipitación decrece rápidamente tierra adentro, de modo que sólo una estrecha faja a lo largo de la costa Norte recibe una cantidad importante de lluvia monzónica. De forma análoga, las áridas tierras centrales del Sahel africano reciben sólo precipitaciones esporádicas durante el verano. Africa occidental y Australia son regiones monzónicas de parecido geográfico, y ninguna de ellas está influenciada por una estructura montañosa dominante que pudiera equipararse al Himalaya.





EL CALENTAMIENTO TOTAL de una columna de aire sobre el continente es la suma de tres contribuciones: calentamiento radiativo (no representado aquí), calentamiento sensible (debido al suelo calentado por el Sol) y calentamiento latente (procedente de la condensación del vapor de agua). La precipitación tiende a reducir el calentamiento sensible de la columna de aire, haciendo que el calentamiento máximo total se traslade tierra adentro. La región donde el aire ascendente alcanza su máxima velocidad también se mueve hacia el interior del continente, y a ello sigue una lenta migración del monzón activo.

Quizás extrañe la ausencia de sistemas monzónicos relevantes en ambas Américas. La región ecuatorial del Hemisferio Occidental está dominada por la cuenca del Amazonas, y la masa de aire situada encima de ésta se halla sometida a inyección de calor sensible y latente durante gran parte del año. En América Septentrional, el enfriamiento radiativo invernal genera considerable energía potencial entre los dos hemisferios. Pero ni siquiera así puede desarrollarse una corriente de aire entre ellos, debido al baluarte opuesto por los Andes.

La fuerza de Coriolis hace que el aire frío de Norteamérica penetre en el océano Pacífico para constituir el alisio del Nordeste. Si la barrera montañosa de los Andes no limitase la circulación, el calentamiento superficial diferencial entre la cuenca del Amazonas y el Pacífico produciría una corriente de aire que penetraría en el Amazonas, dejando una zona de bajas presiones en el Pacífico meridional. Tal zona de bajas presiones obligaría al alisio del Nordeste a cruzar el Ecuador, y la fuerza de Coriolis lo desviaría hacia el Sureste, completando el circuito de corrientes de aire interhemisféricas. Pero la presencia de las montañas determina que la corriente a través del Ecuador sea floja en el Pacífico oriental, y sean los alisios los que continúen soplando a través de todo el Pacífico, hasta Indonesia. Además, el Pacífico oriental está dominado por agua fría. Aun cuando los Andes no existieran, el contenido hídrico del aire resultaría demasiado bajo para crear la alta energía asociada a los monzones asiáticos. El aire que circula en la cuenca amazónica se humedece por la evaporación del océano Atlántico.

Se han realizado cierto número de experimentos de campo a una escala internacional para obtener datos detallados con que investigar los monzones. Merecen citarse el Experimento Internacional del océano Indico, que se llevó a cabo entre 1959 y 1965, y los Experimentos Monzónicos (MONEX), de 1978 y 1979. Los últimos formaban parte del Experimento Meteorológico Global, e incluían estudios separados de las circulaciones monzónicas de verano e invierno. En cada experimento se emplearon satélites, buques y aviones instrumentados para obtener una representación tridimensional del desarrollo del monzón. Una de las aplicaciones más importantes de tales datos será la comprobación de la teorías de las fases activa y durmiente de los monzones.

El programa del alcohol en Brasil

El etanol se está usando en Brasil como combustible puro o mezclado con gasolina. ¿Cuáles son los aspectos técnicos y económicos implicados en su proceso desde la fase agrícola hasta la industrial?

José Roberto Moreira y José Goldemberg

La sustitución de ciertas fuentes de energía por otras más abundantes, o de acceso y uso más convenientes, es un fenómeno bien conocido en la historia de las naciones desarrolladas: la madera, que era la principal (y prácticamente única) fuente de energía usada en gran escala por el hombre hasta el siglo XIX, fue sustituida por el carbón mineral, y éste por el petróleo y el gas natural más recientemente.

Lo que llevó al uso del carbón fue el agotamiento físico de los bosques, que se convirtió en un grave problema en la Inglaterra de la Edad Media. El carbón fue sustituido por el petróleo por motivos de índole tecnológica, dado que los motores de explosión interna (principales consumidores de los derivados del petróleo) exigen combustibles líquidos o gaseosos. Por esta razón, el agotamiento físico de las reservas de petróleo en nuestro siglo (aparte de las dificultades políticas y económicas que conlleva su obtención) no condujo necesariamente al uso del carbón, cuyas reservas están lejos de agotarse.

Al contrario, la "crisis del petróleo" llevó a la necesidad de producir combustibles líquidos a partir del carbón o de la biomasa vegetal. Las tecnologías precisas al respecto existen, como es bien sabido. Durante la segunda guerra mundial, Alemania produjo grandes cantidades de gasolina sintética a partir del carbón, y algunos otros países, entre ellos Brasil, empezaron a usar el vapor de agua, mezclas de gasolina y etanol (y en algunos casos etanol puro) para mover vehículos, reduciendo de este modo el consumo de gasolina.

El problema más grave es el de los costos de producción de estos sustitutivos de la gasolina. Esta es una cuestión muy debatida hoy en los Estados Unidos, si bien fue ya superada en Brasil inicialmente, gracias a una política de subsidios aplicada a partir de 1975. En este país se consumen diariamente cerca de 60.000 barriles de alcohol, y existe en marcha un programa (Programa Nacional do Alcool, PNA) para elevar este número a 170.000 barriles diarios en 1985. En la actualidad, el coste de producción del alcohol compite ya con el de la gasolina, como intentaremos demostrar más adelante.

El uso de etanol en Brasil es ya una realidad, tanto en mezclas del 20 por ciento de etanol en la gasolina, como en alcohol puro (cosa que requiere ciertas alteraciones en los motores). ¿Cuáles son las modalidades de implantación de este programa en Brasil y sus perspectivas en un futuro próximo?

Prácticamente toda la producción de etanol en Brasil se realiza a partir de la caña de azúcar. Por varias razones. La industria del azúcar, extremadamente desarrollada en este país, que es el primer exportador mundial del producto, atravesaba una fase poco favorable en el mercado internacional cuando, en 1975, se propuso el programa nacional del alcohol, es decir, el programa para la utilización del etanol como combustible. Esto despertó inmediatamente el interés de los productores de azúcar, que dedicaron al programa una parte considerable de la capacidad de producción entonces no utilizada en sus instalaciones, desviando una fracción destacada de la masa de caña a la producción de alcohol en vez de azúcar. La tecnología de producción de etanol a partir de caña de azúcar es relativamente sencilla y depende de equipos poco sofisticados, que pueden obtenerse en muchas naciones en vías de desarrollo.

El volumen de inversiones para una fábrica de etanol es relativamente modesto comparado con otras alternativas. El costo tipo de una destilería, con 120.000 litros-día de capacidad, asciende a unos 10 millones de dólares, mientras que la inversión necesaria para una unidad económica de combustible sintético derivado de carbón o pizarras es del orden de 1 billón de dólares (con capacidad para 50.000 barriles) y para una unidad productora de metanol a partir de gas de síntesis asciende a unos centenares de millones de dólares (para una producción de 20.000 barriles-día). A pesar de que el coste de inversión por barril es semejante en estas varias opciones, un programa basado en el etanol obtenido por fermentación de azúcar puede ser introducido lentamente, sin exigir excepcionales inversiones de capital en la fase de viabilización del proceso. Aparte de ello, las destilerías de alcohol pueden encargarse y ponerse en marcha en períodos de sólo doce meses, plazo bastante corto comparado con cualquier otra inversión energética. Estos dos factores resultan especialmente atractivos para naciones en vías de desarrollo, donde la falta de capitales es generalmente un factor dominante en la expansión industrial.

El etanol es un producto bastante común y sus efectos sobre el hombre, muy bien conocidos. El organismo lo tolera en concentraciones atmosféricas muy altas (1000 partes por millón), o sea, dos veces más que la gasolina y, en consecuencia, es muy pequeña la posibilidad de que cause problemas de salud. Siendo un producto orgánico, su impacto ambiental es muy reducido.

Se trata del único sustitutivo de la gasolina que podría fabricarse inmediatamente a escala comercial, ya que fue usado como combustible de automóvil en los primeros años de esta industria, y siguió usándose en motores de coches de competición, en los que se procuraba obtener la máxima potencia.

La forma clásica de preparación del alcohol etílico, a partir de la caña de azúcar, exige una unidad cuyo diagrama de bloque aparece en la ilustración de la página 99. La caña de azúcar es recogida y transportada a destilerías de producción media (como máximo, en Brasil hay unidades de 1 millón de

litros-día de capacidad). En la destilería se lava la caña para la remoción de barro y piedras, luego se tritura y, a continuación, pasa a la unidad de molienda o por un difusor. En la operación de molido atraviesa dos o tres compresores; tras ser de nuevo embebida en agua, pasa por un cuarto compresor. El producto obtenido constituye el vino, que se diluye entonces en fracciones del orden del 10 por ciento de caldo de caña y 90 por ciento de agua.

El vino se almacena entonces en tanques, que se mantienen a una temperatura de 30 a 35 grados Celsius, donde comienza a fermentar. La reacción es exotérmica y hay necesidad de refrigeración, pues a medida que la temperatura rebasa los 40 grados C disminuye el rendimiento bacteriano. La fermentación tiene lugar en un período de 12 a 24 horas. En Brasil se han preferido períodos cortos (12 horas), ya que cuando se alcanza una mezcla con un peso en alcohol del 7 al 8 por ciento la bacteria baja su rendimiento, lo que significa un tiempo prolongado de fermentación para obtener un mínimo aumento en la cantidad de alcohol.

A continuación, la mezcla de alcohol fuertemente hidratado vuelve a ser sometida a destilación y rectificación, y así separar el alcohol. Gracias a la destilación se consigue etanol de una pureza que alcanza los 95,5 grados, pues en este punto se logra una mezcla azeotrópica. Un alcohol aún más puro se obtiene, generalmente, incorporando benceno en la mezcla azeotrópica. Por medio de una nueva destilación (rectificación) se separa el alcohol del benceno y se recupera este último.

Se almacena entonces el alcohol puro. Las exigencias físicas y económicas de esta operación resultan bastante apreciables. Las destilerías de alcohol de caña operan sólo durante una parte del año (de 160 a 180 días), lo que significa la necesidad de tener reservas de este combustible en proporción suficiente para permitir la utilización del producto durante todo el año.

Como se puede ver en el diagrama de bloque, las unidades de procesamiento del etanol cuentan con unidades generadoras de electricidad y vapor, abastecidas por el bagazo de caña que sobra tras la extracción del caldo. Este bagazo, típicamente en una concentración



CAÑA DE AZUCAR. Se ven los colmos compuestos de entrenudos donde se acumula el azúcar. Las hojas se cortan y abandonan en el campo, y sirven como fertilizante. La altura media de las diversas variedades es de 2 metros. En la figura de la derecha se ven las yemas que propagan la especie.

180 GRAMOS DE HEXOSA

92 GRAMOS (64,3 MILILITROS A 15 GRADOS CENTIGRADOS) DE ETANOL

342 GRAMOS DE SACAROSA

360 GRAMOS DE GLUCOSA + FRUCTOSA

184 GRAMOS DE ETANOL

SACAROSA, la principal materia prima del etanol obtenido de la caña de azúcar. Del 14 al 16 por ciento del peso de la caña se debe a la sacarosa. La hidrólisis la descompone en 2 monosacáridos diferentes (glucosa y fructosa). A través de la acción biológica, los monosacáridos se transforman en etanol (y CO_2). Se presenta cantidades estequiométricas. La glucosa existe en la caña en pequeñas cantidades y se transforma en etanol en la proporción mostrada.

de 50 por ciento en peso (es decir, la mitad agua y la otra mitad producto orgánico), está en condiciones de ser quemado y producir vapor.

Hasta ahora se ha venido trabajando con vapor supercalentado de presión bastante modesta, pues lo que se necesita principalmente es un producto en forma de vapor de baja calidad, es decir, vapor que debe transferir calor a la torre de destilación a temperaturas del orden de 100 grados C. Cualquier aumento de presión puede rebajar las necesidades de vapor utilizado para producir trabajo (molienda, transporte de la caña en cintas sinfín, etcétera), aunque esto no presentaría ninguna economía en la operación de destilado. Incluso trabajando con presiones modestas (21 atmósferas), las destilerías producen un exceso de bagazo del orden de 1/3 del volumen total.

En consecuencia, se ha despertado un profundo interés hacia la posible utilización del exceso de bagazo y de presiones más elevadas con objeto de generar electricidad simultáneamente a la producción de vapor (cogeneración). La técnica de cogeneración es bastante conocida y ampliamente aplicada, en Europa sobre todo, aunque en el caso de las fábricas transformadoras de la caña de azúcar se presenta el inconveniente de que sólo funcionan durante medio año.

Existen, no obstante, varios motivos para producir electricidad y venderla a otros usuarios. La primera de ellas es el hecho de que la producción de alcohol coincide con la sequía en los países tropicales, y, en consecuencia, con el período en el que el potencial hidroeléctrico es menor. La segunda es la utilización de otros cultivos (como el sorgo o la mandioca) para permitir extender el período de funcionamiento de la unidad a un plazo mayor; y en esta segunda fase de operación de la fábrica se utilizaría bagazo que se habría almacenado en la época de la zafra y, eventualmente, también madera, si fuera necesario. El almacenamiento de bagazo no presenta ningún problema, y ha sido realizado ya por la industria papelera en Brasil.

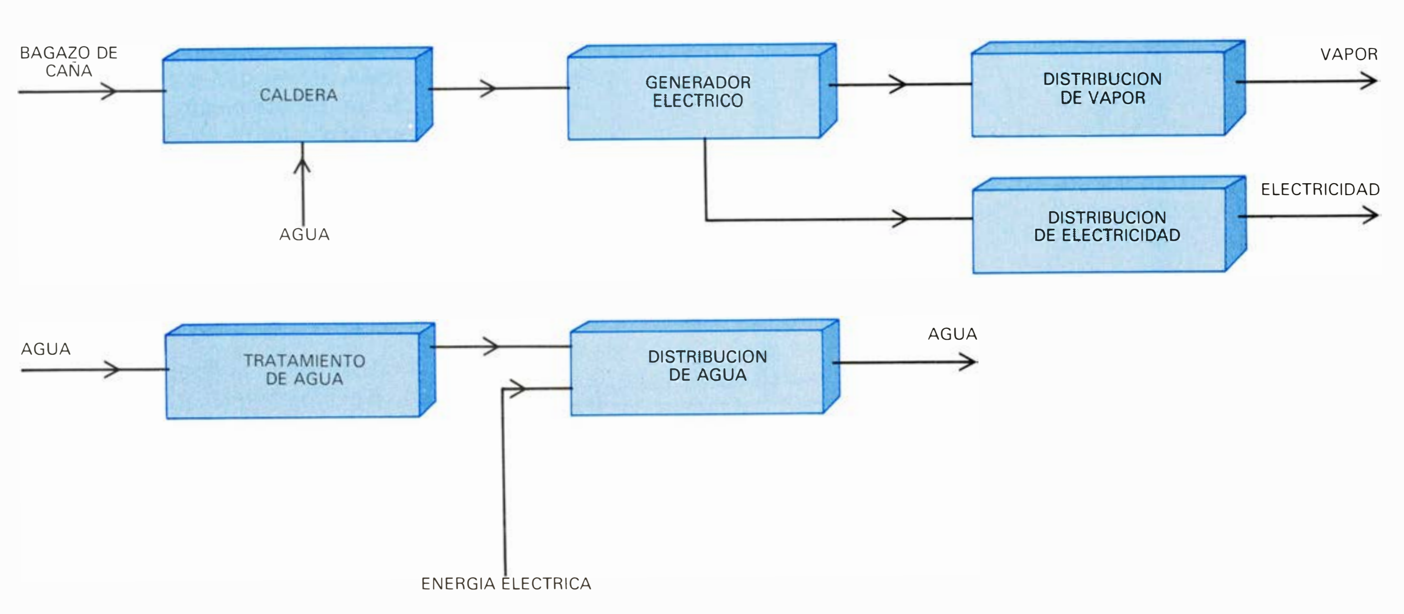
La producción de alcohol de caña en Brasil se ha orientado hacia destilerías autónomas y anexas. Las destilerías anexas fueron construidas junto a las unidades de producción de azúcar, con la finalidad de transferir parte de la materia prima, originalmente usada para producir azúcar, hacia la industria del alcohol. Estas unidades fueron las que se construyeron más rápidamente y con un costo bastante ajustado, pues se trataba de favorecer la expansión de las unidades productoras de alcohol preexistentes (necesariamente una fracción de la melaza se transforma en alcohol) usando instalaciones básicas ya disponibles y tecnología en explotación. Hacia 1977 empezaron a entrar en actividad las destilerías autónomas, así llamadas porque fueron construidas con la única finalidad de producir etanol.

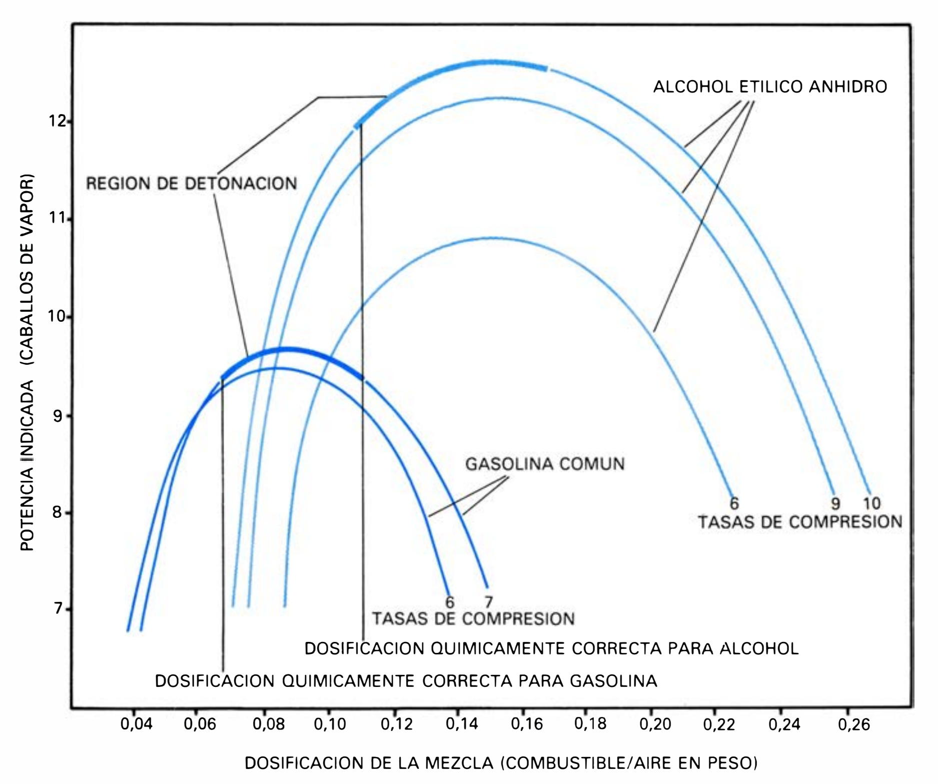
La utilización del etanol como combustible fue afrontada en la fase de introducción del PNA como desventajosa con relación a la gasolina, pues el contenido energético es bastante menor: 5100 kilocalorías por litro para el etanol, mientras que para la gasolina tenemos 7800 kilocalorías por litro. No obstante, si la aplicación del producto es para motores de combustión interna, no es correcto compararlo sólo en relación con el contenido energético, pues hay también otros factores que actúan en favor del alcohol. Como se aprecia en la ilustración superior de la página 100, se pueden usar con alcohol etílico tasas de compresión de hasta 1:12, mientras que en la gasolina están limitadas a 1:8 (en países donde existe gasolina de alta calidad) y 1:6 (en algunos países en vías de desarrollo). Mayor compresión significa necesariamente mayor rendimiento térmico, al menos hasta las tasas de compresión usadas normalmente para motores diesel. A partir de cierto valor, el rozamiento puede resultar tan significativo que no se mejora la eficiencia práctica.

El calor latente de vaporización del alcohol es sustancialmente mayor que el de la gasolina (922 kilojoule por kilogramo) a 25 grados C, mientras que el isooctano (que usaremos como representativo de la gasolina) tiene 307,7 kilojoule por kilogramo. Esto significa que la puesta en marcha en frío con etanol puede presentar dificultades; este problema puede solucionarse usando un calentador eléctrico en la línea de alimentación del combustible, o un tan-

UNA DESTILERIA TIPICA BRASILEÑA usa caña de azúcar como materia prima para el etanol. La caña se corta, se desfibra y se tritura para la extracción del caldo. A éste se adicionan agua y enzimas para producir la fermentación. Los pasos siguientes suponen la recuperación de enzimas, destilación y rectificación (en caso de que se desee etanol prácticamente puro). El alcohol se almacena en recipientes semejantes a los usados para derivados del petróleo. El vapor de agua y la electricidad proporcionan toda la energía necesaria para el proceso, y se generan en calderas de media presión que queman bagazo.

AGUA
MATERIA PRIMA
VAPOR
PICADO Y MOLIDO
AGUA
ENZIMA
VAPOR
FERMENTACION
DIOXIDO DE CARBONO
ENERGIA ELECTRICA
CENTRIFUGACION Y RECUPERACION DEL ENZIMA
VAPOR
DESTILACION
VINAZA
TRANSPORTE Y RETORNO AL SUELO
ALCOHOL 95,5° GL
VAPOR
RECTIFICACION
ALMACENAJE DE ALCOHOL 100° GL





POTENCIA DEL MOTOR en función de la razón combustible-aire y la tasa de compresión, para etanol puro y gasolina. El etanol soporta más tasa de compresión que la gasolina sin sufrir explosión y, asimismo, los motores de alcohol son menos sensibles a variaciones en la razón aire-combustible que los de gasolina.

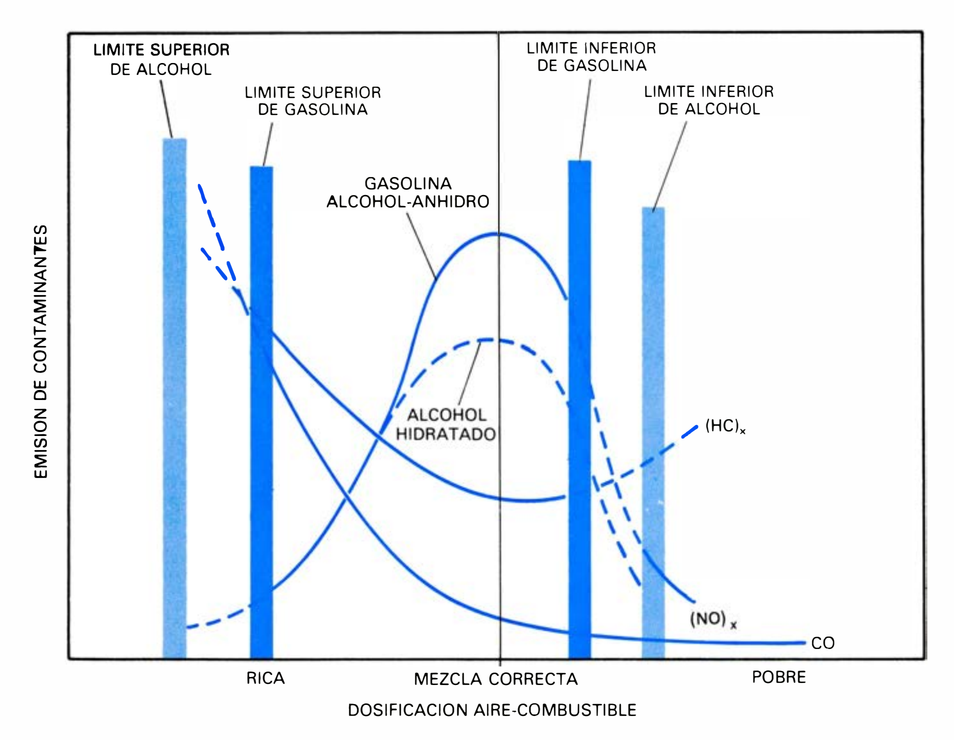
que suplementario de gasolina. Ambos recursos están en uso comercial en Brasil. Esta diferencia de calor latente permite el uso de los gases del escape, aumentando la eficacia del combustible.

Otra posibilidad es evaporar el etanol adiabáticamente, alterando así la densidad de energía en la mezcla de combustible. De hecho, se precisa tres veces más energía para evaporar 1 kilo de etanol que 1 kilo de gasolina. Y, debido al menor poder calorífico, se requiere más masa de etanol que gasolina para un mismo número de kilómetros (o tiempo). Esto significa casi cuatro veces más energía disponible para enfriar el etanol que la gasolina; admitiendo que sea usada la misma cantidad porcentual de energía disponible, puede enfriarse alcohol hasta 0 grados C, mientras que el isooctano se enfriará hasta sólo 22 grados C. En estas condiciones, la mezcla etanol-aire tiene 8,7 por ciento más densidad de energía que la mezcla isooctano-aire. El exceso de peso originado por el volumen extra del tanque del automóvil que va a usar etanol puede llegar al 4 por ciento del peso total del vehículo, y aumentar así el consumo de combustible en un dos por ciento.

No obstante, la mayor potencia disponible en un motor alimentado con etanol debe permitir una reducción en el peso de éste, suficiente para compensar el exceso de peso del combustible. La mayor eficiencia de conversión de energía debe reducir la cantidad total de combustible, valorado sólo en razón del poder calorífico. Todos estos hechos sumados deben convertirse en mejoras de un uno a un 3 por ciento en tasas de compresión de 8,2 y 12, respectivamente.

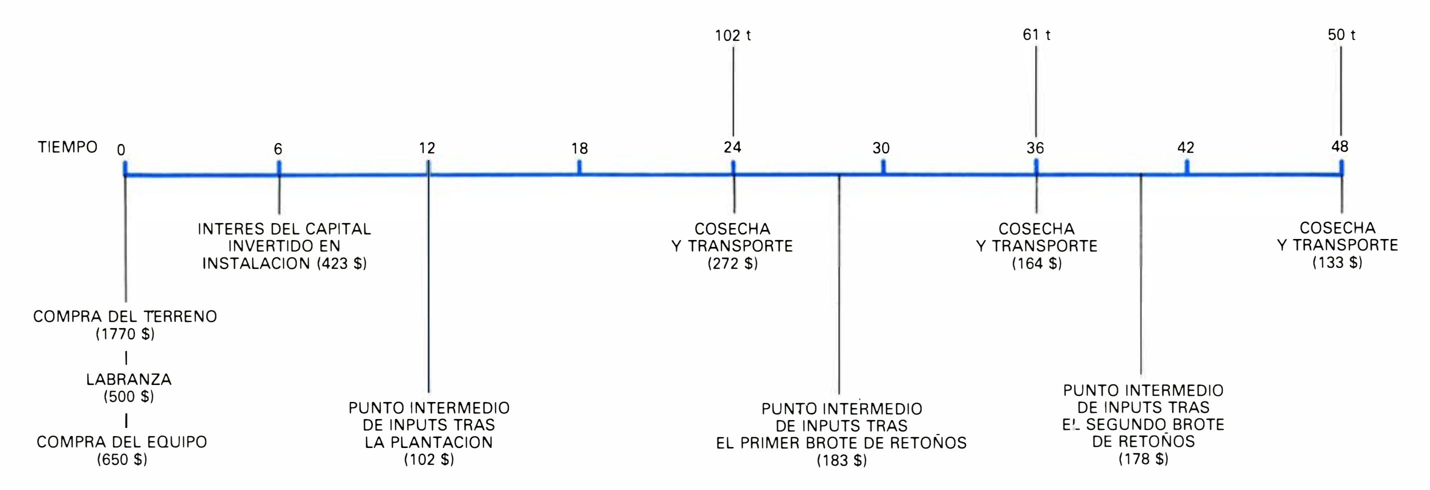
En los vehículos modernos se ha procurado operar con mezclas pobres de combustible, con vista a reducir la contaminación y aumentar el rendimiento. La gran banda de operación de la mezcla aire-etanol [*véase la ilustración inferior de la izquierda*], cuando se compara con la de la gasolina, permite mayor facilidad de operación del motor con mezcla pobre. Ello representa un ahorro de combustible del 2 por ciento.

Podemos, pues, esperar que un litro de etanol, que contiene 5100 kilocalorías de energía, suponga un aumento del 25 por ciento, es decir, un poder calorífico equivalente a 6375 kilocalorías, que se traduce en un exceso de consumo, en comparación con la gasolina, de sólo 7800/6375 = 1,22, o sea, un consumo de un 22 por ciento mayor. Esto concuerda con los resultados que se están obteniendo con los vehículos movidos con alcohol.



LA CONTAMINACION producida por motor tipo Otto [combustible no quemado (HC), monóxido de carbono (CO) y óxidos de nitrógeno (NO_x)] depende de la razón combustible-aire. Como motores de alcohol puro, operan en una franja de mezclas mayor que los de gasolina y hay menos contaminación.

La temperatura de la llama en un motor de combustión interna es de 2440 grados Kelvin para el isooctano y de aproximadamente 2200 grados K para el etanol; como la disociación de los gases en el motor (que da origen a contaminantes como los óxidos de nitrógeno NO_x) depende exponencialmente de la temperatura, en los motores de alcohol se producen cantidades mucho menores de estos contaminantes.



DISTRIBUCION TEMPORAL y costes de las actividades agrícolas para cultivo de la caña. La plantación se hace de cuatro en cuatro años. Dieciocho meses después de la plantación se recoge la caña. Aprovechando el rebrote de la planta (la *soca*) se realizan otras dos cosechas. Como se emplea algún tiempo en la preparación del suelo, el ciclo entero exige 48 meses. El análisis económico del texto supone productividades y costes que aparecen arriba.

Los compuestos orgánicos, hallados en los gases de escape de los motores, son, fundamentalmente, combustible no quemado y aldehídos. Para identificar los factores que afectan a estas emisiones, se han realizado investigaciones por la General Motors y la Volkswagen usando metanol. La principal conclusión a que se ha llegado es la de que, si se ha realizado una cuidadosa preparación del combustible, la cantidad de materia no quemada en los gases de escape puede reducirse entre 80 y 90 por ciento, trazando los niveles de emisión dentro de los patrones aceptables para la gasolina. La cantidad de aldehídos se rebajó también en casi un grado comparada con los sistemas convencionales de alimentación de combustible, situándose en un nivel un poco superior al patrón de emisión de las gasolinas.

La emisión de monóxido de carbono depende principalmente de la razón aire-combustible. En mezclas pobres de combustible, la cantidad de CO es mucho menor que en mezclas ricas. Esto, que depende bastante de la técnica de preparación del combustible, está en relación con el combustible no quemado. Por lo tanto, durante la puesta en marcha de los motores, principalmente en días fríos, cuando es necesaria una mezcla rica en combustible, la cantidad de CO puede ser mayor que la especificada para la gasolina y puede requerir oxidación catalítica para satisfacer patrones tan severos como los que están en vigor en los Estados Unidos.

Existen, pese a todo, muchas otras ventajas en favor del etanol: *a*) eliminación de la emisión de plomo, un contaminante altamente tóxico y presente en las gasolinas convencionales; *b*) reducción de las propiedades carcinogénicas y tóxicas de los compuestos aromáticos encontrados en la gasolina; *c*) ausencia de compuestos de azufre.

Se necesitarán estudios más profundos para establecer una comparación final entre la contaminación del etanol y la de la gasolina. Entretanto, cuando se habla de mezclas de 10 a 20 por ciento de alcohol con gasolina, pocos cambios podrán esperarse en el patrón de la contaminación obtenido de las gasolinas, pues los resultados experimentales disponibles nunca muestran un aumento de contaminantes particulares del alcohol de cierta magnitud, siempre que se adopten los cuidados convenientes en la preparación del combustible. Desgraciadamente, aún no se dispone de mediciones extensas de los patrones de contaminación en automóviles que usen sólo etanol.

La producción de etanol a partir de la caña de azúcar (o de otros productos agrícolas) exige cantidades apreciables de energía, tanto en la fase agrícola como en la fase industrial. En el caso de que esta energía derive de combustibles fósiles, y de petróleo en particular (caso frecuente), importa analizar este problema cuidadosamente, dado que es posible, en principio, que la producción de un litro de etanol (con 5100 kilocalorías) exija una cantidad mayor de energía bajo la forma de derivados del petróleo.

En el caso de la caña de azúcar, la situación es particularmente favorable, porque, aparte de los azúcares, la caña produce bagazo en cantidad suficiente para abastecer de toda la energía necesaria en la fase industrial del proceso.

Para la evaluación de la cantidad de etanol que puede obtenerse, se ha de considerar que la eficiencia de extracción de los azúcares monosacáridos y disacáridos del colmo de la caña es del 95 por ciento y el proceso de fermentación presenta también una eficiencia del 95 por ciento. Esto significa que una tonelada de colmo de caña, con composición media representada en la ilustración central de la página 102, puede producir unos 90 litros de etanol. En la práctica, la producción típica en Brasil es de 66 litros/tonelada de caña y 250 kilogramos de bagazo (con 50 por ciento de humedad), o sea, cerca de 4 kilogramos de bagazo por litro de etanol. Un kilogramo de este bagazo tiene un poder calorífico de 1800 kilocalorías y puede producir 2,44 kilos de vapor (1 kilo de vapor necesita 900 kilocalorías), cuando se considera la eficiencia de las calderas donde se quema. La productividad típica de la caña en el sur del Brasil es de cerca de 54 toneladas por hectárea y año, con una productividad total de 3500 litros por hectárea y año.

La energía consumida en la preparación del etanol se debe, en parte, a la fase agrícola (plantado y recolección de la caña) y, en parte, a la fase industrial (extracción del caldo de caña, fermentación y destilación). Se computa siempre el consumo directo e indirecto de la energía; verbigracia: la energía necesaria para producir fertilizantes o maquinaria agrícola.

Como resultado de este tipo de análisis, se concluye que el 85 por ciento de energía gastada en la fase agrícola lo es directamente en forma de combustibles (48 por ciento), o en combustibles usados en la fabricación de las propias máquinas (15 por ciento) y en la producción del nitrógeno usado como fertilizante (22 por ciento); el resto de la energía consumida lo es en trabajo, preparación de las simientes, herbicidas, etcétera. En la fase industrial, el

	TASA DE COMPRESION 8,2:1	TASA DE COMPRESION 12:1
1. Tasa de compresión	0	+ 16
2. Eficiencia debida a inyección de combustible	+ 8	+ 8
3. Exceso de potencia	+ 1	+ 3
4. Exceso del peso del combustible	− 2	− 2
5. Operación con mezcla pobre	+ 2	+ 2
TOTAL	+ 9%	+ 25%

LA EFICIENCIA DE UN VEHICULO DE ALCOHOL es muy superior a lo que se podría esperar basándose en el poder calorífico del combustible. El aumento de la tasa de compresión de 8,2:1 a 12:1 significa un aumento del 16 por ciento, en lo que se refiere a la potencia. Otras innovaciones usadas también en los motores de gasolina suponen un aumento de potencia del 25 por ciento. (Tabla de los autores del artículo.)

COMPONENTE	% DEL COLMO POR PESO
Sacarosa	12-16
Azúcares reductores (glucosa y fructosa)	0,2-1,5
Azúcares totales (en porcentaje de glucosa)	13-17
Fibra	9-13
Humedad	70-79

COMPONENTES MAYORITARIOS de la caña de azúcar. La humedad es grande y limita la explotación agrícola a áreas próximas a la destilería (20-30 kilómetros). La fibra es un componente útil, pues se usa como combustible en las calderas, aportando el vapor y la electricidad necesarios en la industrial.

90 por ciento de toda la energía se aprovecha como combustible para producir vapor. La fase industrial consume casi 3 veces más energía que la fase agrícola.

La tabla de la ilustración superior de la página siguiente compara la energía producida por litro de etanol (5100 kilocalorías) con la energía necesaria para obtenerlo, utilizando dos hipótesis: *a*) toda la energía consumida procede de combustibles fósiles y *b*) toda la fase industrial emplea como fuente de energía el propio bagazo; como hemos visto antes, cada litro de etanol produce 4 kilos de bagazo, o sea, 7200 kilocalorías; de los cuales 2/3 bastan para cubrir los gastos en la fase industrial. Como se aprecia en la tabla, hay siempre una ganancia de energía en la producción de etanol de caña, es decir, la energía invertida en la caña de azúcar es *amplificada* por el proceso de plantado, crecimiento y extracción del etanol (evidentemente, a costa de la energía solar responsable de la fotosíntesis).

Conforme se ha mencionado ya en la introducción, inmediatamente después de cuadruplicarse el precio del petróleo, a finales de 1974, el gobierno brasileño propuso un programa de sustitución de los derivados del petróleo, que se habría de llevar a cabo en varias etapas. En la segunda ilustración de la página siguiente se señalan los objetivos determinados en aquella fecha; a la primera etapa se le atribuyó un plazo máximo, es decir, que se decidió que habría de ser posible adicionar un 20 por ciento de alcohol a toda la gasolina vendida en el país en el año 1980. La adición del alcohol a la gasolina se estaba realizando, desde 1950, en fracciones que llegaron en algunas épocas a niveles del 8 por ciento y, en consecuencia, parecía razonablemente seguro, desde el punto de vista técnico, que pudieran operar con este nivel de alcohol, sin modificaciones, los motores proyectados para gasolina, aunque disminuyera su rendimiento total. La segunda fase, es decir, la de sustitución total de la gasolina por el alcohol, suponía estudios y cambios técnicos en los motores para obtener un buen rendimiento. Aparte de ello, se planteaban serios problemas económicos, pues las refinerías de petróleo se habían proyectado para atender a un mercado que exigía cantidades prácticamente iguales de gasolina, diesel y aceite combustible. De este modo, la reducción excesiva de la demanda de gasolina no podría ser atendida por la industria del refinado de petróleo sin perjudicar el mercado del diesel y del aceite combustible. La tercera y cuarta fases presentaban dificultades aún mayores, pues aparte de exigir una alteración en la industria del refinado, se tendrían que vencer obstáculos técnicos difíciles de evaluar entonces, ya que nunca se había usado alcohol en motores diesel.

FASE AGRICOLA	FASE INDUSTRIAL	TOTAL
1150	3000	4200

ENERGIA NECESARIA para la formación de etanol de caña de azúcar, en kilocalorías por litro.

Para incentivar la producción de etanol se preparó un amplio programa. El gobierno proporcionaba el 80 por ciento del capital (y en algunas regiones el 90 por ciento); la iniciativa privada aportaba el resto. El dinero del gobierno se pagaría en un plazo de cinco años con intereses negativos, es decir, con intereses y corrección monetaria razonablemente por bajo del índice oficial de desvalorización de la moneda. Por otra parte, existía ya en el país una estructura industrial capaz de atender la demanda de nuevas destilerías. Hasta el momento se han autorizado por el gobierno un total que excede las 250 nuevas unidades, y hay ya cerca de 200 en producción. La unidad más común es la que tiene una capacidad de 120.000 litros por día, con un coste aproximado de 10 millones de dólares. En 1980, la producción de etanol alcanzó el objetivo determinado en 1974 (4 millones de litros por año), y se estima que poco más de tres billones serán producidos por las unidades instaladas dentro del Programa Nacional do Alcool, lo que significa una inversión aproximada de 2 billones de dólares.

Aparte de la subvención financiera de capital, el alcohol recibió apoyos a través de la eliminación de las tasas existentes para la gasolina, responsables de casi el 30 por ciento del precio de venta al consumidor. Conviene añadir, por último, que la gasolina siempre ha tenido un coste artificialmente elevado, para permitir precios más ventajosos al diesel (usado para fines comerciales) y al aceite combustible (usado para fines industriales). Con todas estas ventajas económicas, el alcohol se vendió, a partir de 1975, a un precio que resultaba inferior en un 40 por ciento al de la gasolina por litro.

En 1979, el éxito del PNA resultó tan evidente, gracias a la política económica adoptada y a la constante elevación del precio del petróleo, que el gobierno decidió adoptar un plazo para la ejecución de una segunda fase del programa, aunque un poco menos ambiciosa que la establecida en 1975. Se estableció un acuerdo entre los productores de automóviles y el gobierno, dirigido a la producción en estos tres años (80-82) de

900.000 nuevos automóviles totalmente movidos con etanol, mientras que otros 280.000 serán modificados a fin de que puedan utilizar también este combustible. El gobierno garantizó una producción de 10,7 billones de litros (210.000 barriles/día) para el año 1985, poniendo a disposición de la iniciativa privada un capital que ascendía a 5 billones de dólares. Hasta el momento han sido sometidas a comprobación varias flotillas de automóviles, con un total de más de dos mil unidades, y se ha llegado a las siguientes conclusiones. En primer lugar, las modificaciones en los motores, bastante sencillas, suponen fundamentalmente un aumento en la tasa de compresión de 1:6 a 1:12; una alteración en el carburador, que exige mayor cantidad de alcohol para la misma cantidad de aire que los automóviles tradicionales; la adición de un sistema auxiliar para facilitar la primera ignición en días de temperatura por debajo de 10 grados C. (Kits para modificaciones del motor a gasolina para operaciones con alcohol han sido puestos ya a la venta a precios relativamente modestos.) Y, en segundo lugar, el consumo de alcohol es un 25 por ciento superior al de gasolina, incluso después del aumento de la tasa de compresión.

La fase siguiente del programa estriba en la utilización de etanol en motores diesel, con una simultánea sustitución del aceite combustible industrial por otro producto. Se ha investigado intensamente el uso de mezclas diesel-etanol en los últimos cuatro años y se han conseguido operaciones hasta con 70 por ciento de etanol. La mayor dificultad de la operación está ligada a la alta resistencia presentada por el alcohol a la ignición por compresión, lo que se mide técnicamente por el índice de cetano atribuido al producto. Una de las propuestas para usar etanol en esta aplicación es la elevación de su índice de cetano a través de la adición de productos químicos con características de explosivos, el amino por ejemplo.

Otra posibilidad, menos dependiente de las investigaciones que se están realizando, sería la adición de bujías de ignición en los actuales motores diesel. Para la sustitución del aceite combustible no tiene Brasil aún una política definida, pero existe la posibilidad de utilizar carbón, madera natural, madera peletizada y particulados de madera que pueden ser fácilmente adicionados al combustible.

Al iniciarse la implantación del PNA, no había dudas de que el etanol no podría competir en precio con la gasolina. Pero desde entonces, la importación de grandes cantidades de petróleo (Brasil importa el 85 por ciento del petróleo que consume) constituye una carga excesiva para la balanza de pagos del país, y la coloca en situación crónicamente deficitaria. En este contexto es preciso recordar que la energía tiene dos niveles distintos de precio: uno es el precio que el consumidor puede pagar y otro el que el país puede pagar. En los últimos años, principalmente en 1979, el petróleo alcanzó precios muy elevados y actualmente se negocia a un nivel de 35 dólares el barril. En estas condiciones, el alcohol presenta condiciones competitivas mejores, como veremos a continuación.

Para hacer un análisis de los costes reales, usaremos tres tasas distintas de interés sobre los préstamos necesarios para financiar las inversiones que se precisan para la producción de alcohol. Estas financiaciones son, en Brasil, gubernamentales. Una tasa mínima del 3,6 por ciento resulta de la suposición de que la financiación del gobierno, sobre el 80 por ciento de la inversión, se realiza a la tasa de 3 por ciento anual y el resto lo financia la iniciativa privada, a un valor del 6 por ciento. La tasa del 3

HIPOTESIS	ENERGIA PRODUCIDA POR LITRO DE ALCOHOL-A (KCAL)	ENERGIA TOTAL NECESARIA-B (KCAL)	BALANCE ENERGETICO A/B
Toda la energía necesaria procede de derivados del petróleo	5100	4200	1,2
Toda la energía de fase industrial procede del propio bagazo	5100	1150	4,4

EL BALANCE ENERGETICO depende de las fronteras escogidas para la actividad agrícola e industrial y del grado de desarrollo del país [*véase la ilustración precedente*]. Se recogen aquí los resultados para dos fronteras diferentes. Los resultados son válidos sólo para prácticas agrícolas semejantes a las brasileñas.

	PRODUCCION (LITROS/AÑO) X 10^9	AREA CULTIVADA DE CAÑA DE AZUCAR EN 1000 HA
Hipótesis I	3	1100
Hipótesis II	16	4400
Hipótesis III	22	6000
Hipótesis IV	33	9000

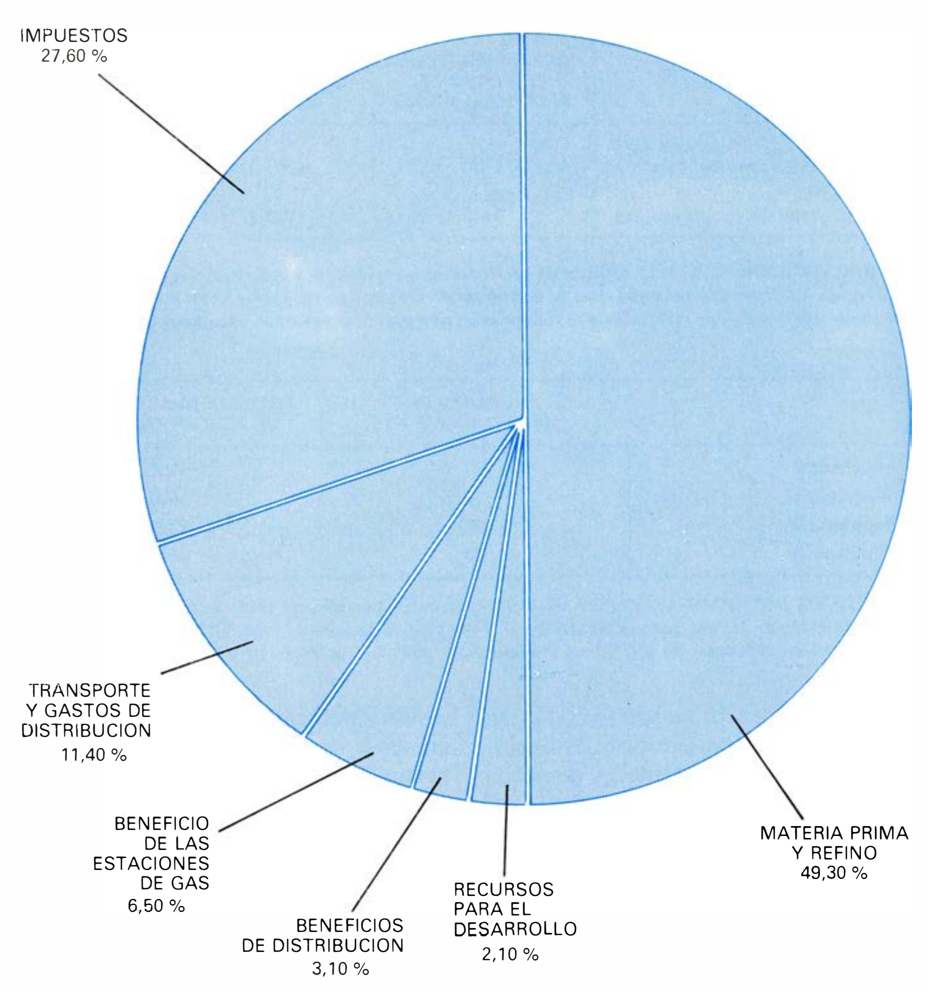
OBJETIVOS DETERMINADOS POR EL PNA cuando fue lanzado, en 1975. Se trataba de utilizar el etanol extraído de la caña para combustible. El Tope I fue alcanzado en 1980. El Tope II, redefinido en 1979, prevé la producción de 10,7 billones de litros/año en 1985. Los Topes III y IV no tienen aún plazos.

SECTOR	US$/M³ DE ETANOL (21 GJ)
Enzimas, productos químicos y energía	6,00
Mano de obra	15,00
Materiales de manutención, operación, seguro y administración	25,00
TOTAL	46,00

COSTES OPERATIVOS DE LA FASE INDUSTRIAL en la conversión de la caña de azúcar en etanol. Las cifras recogidas aquí corresponden a una destilería que rinde 120.000 litros diarios de etanol.

SECTOR	US$/GJ (1980)
Inversiones fijas en la destilería	1,90
Operación y mantenimiento	2,20
Materia prima	9,23
Crédito del sub-proyecto	(0,70)
Capital de giro para la operación	0,05
	12,68
Inventario del proyecto	0,46
	13,14

ANALISIS ECONOMICO PARA LA UNIDAD de 120.000 litros de etanol/día, tasa de interés 6 por ciento. El coste de la materia prima es aspecto dominante en el precio del etanol. Importa, por tanto, aumentar la productividad agrícola para mejorar la competitividad del etanol frente a la gasolina.



COMPOSICION DEL PRECIO DE LA GASOLINA en Brasil. Casi la mitad de la inversión se la llevan los gastos de materia prima y refino. Brasil importa el 85 por ciento del petróleo que consume, lo que constituye una carga para su balanza de pagos y la coloca en una situación crónicamente deficitaria.

por ciento se eligió por ser éste el coste del dinero en dólares constantes en los Estados Unidos. En el otro extremo, usaremos una tasa del 12 por ciento, sugerida por algunos economistas para países en vías de desarrollo. Entre ambos extremos, elegiremos la tasa del 6 por ciento como caso base.

La magnitud de las inversiones y de la producción agrícola aparece detallada en la ilustración de la página 101 y es el resultado medio para la zafra del 76/77 en la región sudeste del Brasil. La caña tiene una productividad de 54 toneladas anuales, con una humedad del 75 por ciento y un poder calorífico de biomasa de 16,7 gigajoule por tonelada; así pues, una hectárea de caña de azúcar posee un contenido energético de 226 gigajoule. Podemos entonces calcular el precio de la materia prima en 2,38 o en 3,95 dólares por gigajoule, según se englobe o no el coste de la tierra y de la tasa de intereses utilizada. La inclusión del coste de la tierra es un problema complejo, y si hay que atribuirle algún interés, la tasa más realista parece el 3 por ciento; con esta tasa de interés la tierra representa sólo el 10 por ciento de los gastos agrícolas totales.

El coste del capital de la destilería tiene dos componentes mayoritarios: la inversión fija y el capital circulante. Este comprende, entre otros costes, el inventario de la materia prima y del producto final. En el caso de la industria del alcohol, el inventario del producto final es muy importante, pues la industria opera sólo durante el período de la cosecha, o sea, de 160 a 180 días. Entonces hay que almacenar una cantidad razonable de producto para atender a la demanda durante todo el año. Calculando 20 años como vida media de la destilería, y un valor de 3600 dólares por gigajoule y día de capacidad instalada, obtenemos un valor para la inversión fija de 1,90 dólares por gigajoule, para una tasa de intereses del 6 por ciento.

La única biomasa necesaria para la producción de alcohol es la caña de azúcar, pues el bagazo es más adecuado para proporcionar la energía requerida en el procesado de la caña. Como hemos visto, una tonelada de caña, con un contenido energético de 226 gigajoule, produce 3500 litros de alcohol con un contenido energético de (3500 × 5100 kilocalorías) × 74,6 gigajoule, o sea, un rendimiento del 33,0 por ciento. Esto significa, pues, un coste de materia prima entre 7,21 y 11,97 dólares por gigajoule.

El único producto suplementario derivado de la caña de azúcar es la vinaza, que se usa como fertilizante en la mayoría de los centros de cultivo de caña del Brasil. Los excedentes de bagazo, a pesar de que éste sea extrínsecamente valioso, no son considerados de valor económico, ya que hasta ahora no se usan. Un criterio típico para la sustitución del fertilizante por la vinaza es el de 0,70 dólares por gigajoule de alcohol.

A los costes evaluados ya, hay que sumar el valor del capital de giro. El valor típico es del 5 por ciento de la inversión fija más la contribución de la inversión vinculada al almacenamiento de alcohol. Suponiendo un valor de 12,68 dólares por gigajoule, resultante de los costes anteriores, y admitiendo que el 6 por ciento de la producción va dirigido al inventario del producto (evidentemente, esto no ocurre durante todo el año, pues una vez establecido el inventario, éste permanece constante) concluimos que éste significa 0,6×165× × 12,17=1255 dólares. Se trata de una inversión no despreciable; su coste anual se computa multiplicando este valor por la tasa de intereses. Esto significa 0,46 dólares por gigajoule. Lo que demuestra que el coste total del etanol es de 13.14 dólares por gigajoule.

Considerando que la producción de un gigajoule de gasolina de buena calidad requiere aproximadamente 1,22 gigajoule de aceite crudo, y que un barril de crudo contiene (8250 kilocalorías por 159 litros = 5,48 gigajoule), podemos producir un máximo de 4,42 gigajoule de gasolina, lo que significa un valor mínimo de 7,92 dólares por gigajoule, como contribución de la materia prima a la gasolina. A este coste, debemos añadir los gastos de operatividad, manutención, amortización de capital y mano de obra, conforme hemos hecho para el alcohol. El valor típico puede obtenerse multiplicando el valor del gigajoule de petróleo por 1,64, o sea, 10,47 dólares por gigajoule.

Como se ha visto antes, la eficiencia del etanol es un 25 por ciento superior a la de la gasolina, lo que muestra que, en términos de coste por kilómetro recorrido, el etanol producido en Brasil compite ya con la gasolina, independientemente de cualquier subvención.



RETRATO DE SADI CARNOT pintado por Louis Léopold Boilly en 1813, cuando Carnot tenía 17 años. Viste el uniforme de la École Polytechnique de París, centro fundado en 1794 para formar ingenieros militares; al emprender su trabajo sobre las máquinas de vapor tenía ya una breve carrera castrense. Siendo aún estudiante, le pidió a Napoleón que permitiera a los alumnos de la École Polytechnique participar en la defensa del imperio, atacado por varios ejércitos europeos. Napoleón accedió y, cuando París fue asediada en marzo de 1814, Carnot y sus compañeros lucharon en las afueras de la ciudad.

Sadi Carnot

Aunque se le conoce por su análisis de una máquina térmica ideal, su interés estaba en la aplicación práctica de la potencia mecánica del vapor. Más aún, el ciclo que describió no es el ideal para la mayoría de las máquinas modernas

S. S. Wilson

Nicolas Léonard Sadi Carnot fue un miembro destacado de una familia extraordinaria. Su padre, Lazare Nicolas Marguerite Carnot, fue un distinguido general que sirvió como ministro de la guerra bajo las órdenes de Napoleón y escribió tratados de estrategia militar, mecánica, geometría y fundamentos del cálculo. El hermano más joven de Sadi, Lazare Hippolyte, fue periodista y político radical. El sobrino de Sadi, Marie François Sadi Carnot, presidió la República Francesa desde 1887 hasta 1894, muriendo este último año a manos de un anarquista italiano. El propio Sadi Carnot entró en la historia como el filósofo de la naturaleza (lo que ahora llamaríamos un físico) que en 1824 fundó la ciencia de la termodinámica con la obra *Réflexions sur la puissance motrice du feu et sur les machines propres à développer cette puissance (Reflexiones sobre la potencia motriz del calor y sobre las máquinas adecuadas para desarrollar esta potencia)*. Se le atribuye, en particular, la definición de un ciclo ideal de funcionamiento para una máquina térmica, es decir, un ciclo con el rendimiento máximo posible. El rendimiento térmico es el cociente entre el trabajo producido y el calor consumido; se expresa en función de la temperatura absoluta de la fuente de calor y la temperatura absoluta del componente de la máquina que absorbe el calor.

Esta interpretación de los aspectos esenciales de la obra de Carnot puede inducir a engaño, ya que las preocupaciones de este autor no se inscribieron en el dominio que hoy consideramos propio de un físico. No concibió su análisis de las máquinas térmicas como un tratado teórico, sino como una obra de divulgación destinada a mejorar y extender la utilización en Francia de las máquinas de vapor y otros ingenios térmicos. El ciclo que Carnot describe no es un ciclo ideal para el funcionamiento de las máquinas térmicas reales. Tampoco definió el rendimiento térmico del ciclo, porque no acabó de entender del todo la equivalencia entre calor y trabajo como formas de energía y porque el concepto de temperatura absoluta le era desconocido.

La obra de Carnot ha sido entendida mal, principalmente porque pocos la han leído. Publicadas por Bachelier, el principal editor de trabajos científicos de Francia, las *Réflexions* fueron objeto de reseñas favorables y la edición se agotó pronto. Pero hasta 1890 no se tradujo al inglés. Hoy, muchos ingenieros y científicos están familiarizados con la obra de Carnot, no a través de sus propios escritos, sino a través de los comentarios de mediados del siglo XIX debidos a Benoît P. Émile Clapeyron, Rudolf Julius, Emanuel Clausius y William Thomson (convertido después en Lord Kelvin). Fue Thomson quien expresó el rendimiento térmico del ciclo hipotético de Carnot en términos de la temperatura absoluta. En los comentarios mencionados se puso excesivo énfasis en el papel del ciclo de Carnot como un ciclo ideal para las máquinas térmicas, con menoscabo del resto de la obra de Carnot, que está llena de recomendaciones prácticas bien fundamentadas. Pienso que si sus contemporáneos hubiesen leído y seguido estas recomendaciones, la revolución industrial en Francia se habría acelerado.

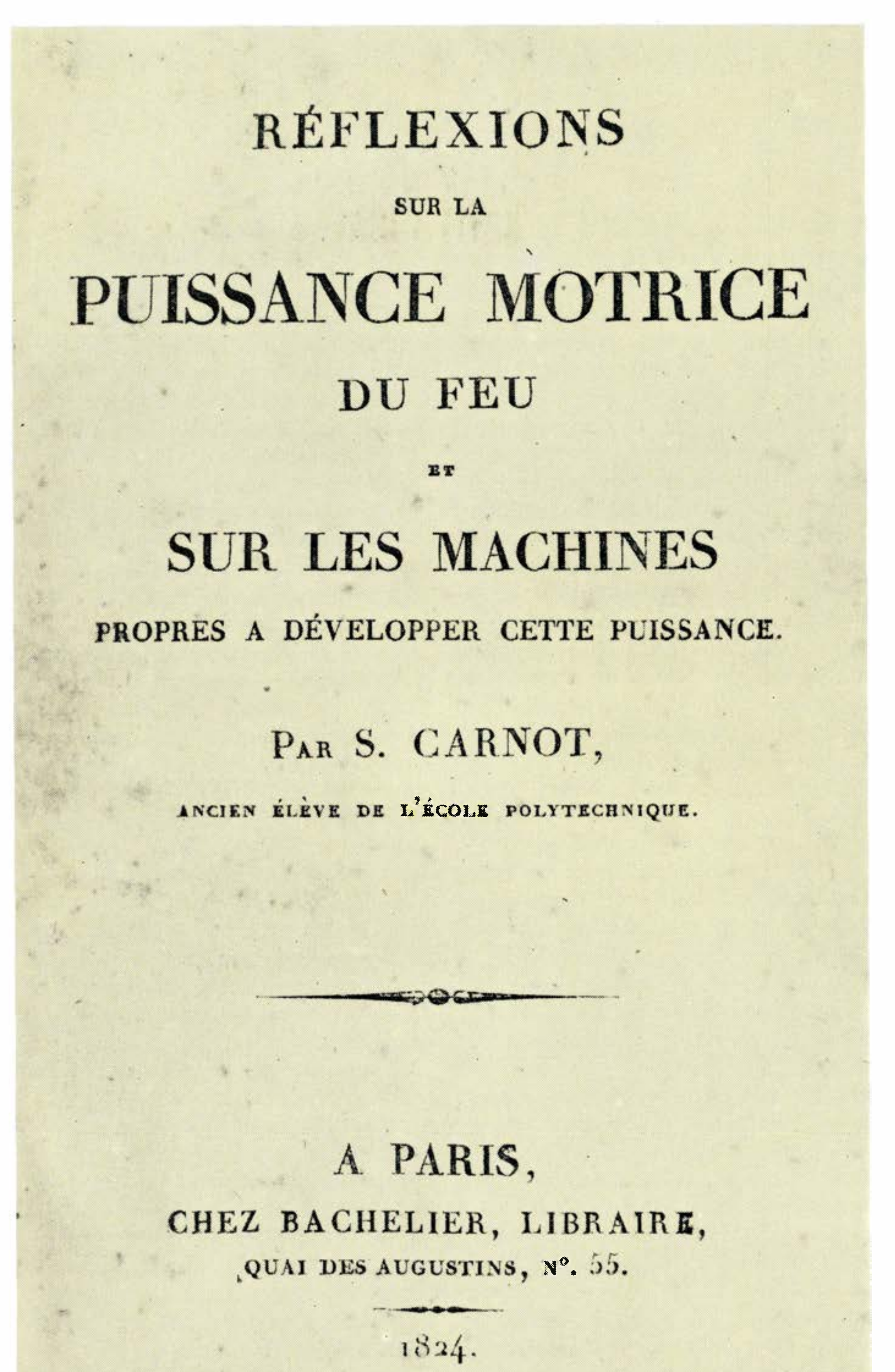
Carnot nació en 1796, hacia el final del período de la historia europea que ha sido denominado la Edad de la Razón. Dos años antes se había fundado la École Polytechnique para formar ingenieros militares. En los primeros 35 años de existencia de esta institución, pasaron por ella y allí se educaron una serie de destacados autores en las ciencias físicas y matemáticas, entre los que figuran Joseph Louis Lagrange, Jean Baptiste Joseph Fourier, Claude Louis Berthollet, André Marie Ampère, Pierre Louis Dulong, Augustin Louis Cauchy, Charles Bernard Desormes, Nicolas Clément-Desormes, Gaspard Gustave de Coriolis, Siméon Denis Poisson, Joseph Louis Gay-Lussac, Augustin Jean Fresnel, Jean Baptiste Biot, Jean Louis Marie Poiseuille y Clapeyron. Carnot también se formó en la École Polytechnique.

El trabajo de esa pléyade contribuyó a desarrollar la creencia en que todos los aspectos del mundo físico podían entenderse a la luz de la razón, siempre que se hiciera el esfuerzo suficiente para aplicar el método científico del análisis teórico, sustentado en la experimentación. El método científico triunfó admirablemente en mecánica, electricidad, óptica y en el análisis de materiales, sirviendo también de fundamento a la ciencia de la ingeniería, que desplazó el enfoque empírico de la tecnología que había resultado fructífero hasta los comienzos del siglo XIX. En el siglo XVIII, particularmente en Gran Bretaña, los expertos constructores de máquinas e instrumentistas habían realizado la revolución industrial, no a través del análisis teórico, sino mediante reglas empíricas que condujeron a mejoras en agricultura, en la utilización de la energía del viento, agua y vapor, en la construcción de canales, ferrocarriles y barcos de vapor, en la fundición del hierro con carbón y en la mecanización de la producción textil y otras industrias. La revolución industrial tuvo lugar en Gran Bretaña mucho antes que en Francia, el resto de Europa y los Estados Unidos.

El retraso de la industrialización de Francia se explica en parte por la inestabilidad política que resultaba del vacilante poder de la monarquía y la aristocracia, el ascenso y caída de Napoleón y la disolución del imperio francés tras la derrota de Napoleón en Waterloo. Pero los factores tecnológicos

también contribuyeron a ese retraso. La industria carbonífera francesa distaba mucho de la pujanza británica. Y no construyeron su primer ferrocarril de pasajeros hasta la década de 1840, e incluso entonces la obra fue supervisada por un contratista británico, Thomas Brassey.

Sadi Carnot había nacido en París, en el Palais du Petit Luxembourg, donde su padre residía por ser miembro del Directorio que gobernó Francia entre el final de la revolución, en 1795, y la subida de Napoleón al poder, en 1799. Lazare se retiró de la vida pública en 1807 para enseñarles a sus hijos Sadi e Hippolyte matemáticas, física, lenguas y música. Inscribió al primero en la École Polytechnique en 1812, cuando contaba 16 años, la edad mínima para la admisión.


RÉFLEXIONS

SUR LA

PUISSANCE MOTRICE

DU FEU

ET

SUR LES MACHINES

PROPRES A DÉVELOPPER CETTE PUISSANCE.

PAR S. CARNOT,

ANCIEN ÉLÈVE DE L'ÉCOLE POLYTECHNIQUE.

A PARIS,

CHEZ BACHELIER, LIBRAIRE,

QUAI DES AUGUSTINS, N°. 55.

1824.

EL PRINCIPAL TRABAJO DE CARNOT, publicado en 1824, promovió la construcción de máquinas de vapor y otros ingenios térmicos en Francia, cuyo desarrollo industrial iba rezagado con respecto a Inglaterra. Aquí se muestra la portada. El título, traducido, se lee *Reflexiones sobre la potencia motriz del calor y sobre las máquinas idóneas para desarrollar esta potencia*. En las *Réflexions*, Carnot definió un ciclo hipotético de funcionamiento para una máquina térmica que podría generar mayor potencia motriz que cualquier otro ciclo. La naturaleza del ciclo se interpretó equivocadamente entre la comunidad científica.

El año siguiente, los ejércitos aliados invadieron Francia y Sadi escribió a Napoleón pidiéndole que permitiera a los estudiantes incorporarse a la defensa del imperio: "Señor: El país necesita a todos sus defensores. Los alumnos de la École Polytechnique, fieles a su lema, solicitan les sea permitido dirigirse con prontitud a las fronteras para compartir la gloria de los bravos hombres que están entregando su vida a la salvación de Francia. El batallón, orgulloso de haber contribuido a la derrota del enemigo, volverá a la escuela a cultivar las ciencias y a prepararse para nuevos servicios." Napoleón concedió el favor pedido y, en marzo de 1814, los estudiantes combatieron con bravura, aunque sin éxito, para impedir la entrada en París a los ejércitos aliados.

Napoleón abdicó en abril, pero volvió al poder un año después. En el período subsiguiente, conocido como los Cien Días, Lazare Carnot ocupó la cartera ministerial del interior. Después de que Napoleón abdicase por segunda vez, tras su derrota en Waterloo, Lazare se exilió. Sadi, que se había graduado y servía en el ejército como oficial ingeniero, fue relegado a trabajos de rutina en la guarnición. En 1820 se retiró del ejército con media paga y regresó a París. Allí estudió física y economía política en la Sorbonne, el Collège de France, la École des Mines y el Conservatoire des Arts et Métiers. También visitó diversas industrias y talleres para observar cómo operaban, en la práctica, los principios de la economía.

Del carácter de Carnot no se sabe mucho. Se le describió como un hombre reservado, reacio a la publicidad; se le atribuye la frase: "Dí poco acerca de lo que sabes y nada sobre lo que no sabes". La mayor parte de la información sobre su vida procede de un apunte que Hyppolite escribió en 1878, unos 46 años después de la muerte de Sadi. Señala su hermano que el nombre de Sadi procede de un poeta y moralista persa, Sa'dī o Shīrāz, cuyos escritos gustaban a Lazare. Según el mismo Hippolyte, Sadi fue de constitución delicada, fortalecida a través del ejercicio físico. Estaba dotado de una sensibilidad extremada y de una gran energía. Cuando consideraba que estaba combatiendo la injusticia, nada podía detenerlo.

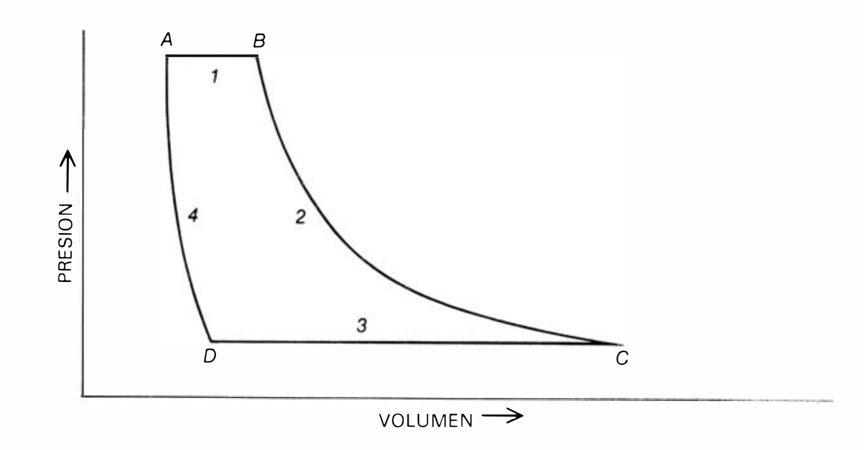
No está claro cómo llegó a interesarse Carnot en las máquinas de vapor. Robert Fox, de la Universidad de Lancaster, el principal estudioso de la obra de Carnot, afirma que encontró su inspiración, no en la física de su época, sino en la tradición procedente de la ingeniería. Carnot mantuvo largas discusiones con los químicos e industriales Charles Bernard Desormes y Nicolas Clément-Desormes, quienes celebraron conjuntamente una conferencia en 1819 en torno a la teoría del uso de la expansión del vapor, una tecnología común que James Watt había patentado en Inglaterra. El uso de la expansión del vapor se discutió mucho por esa época en Francia, debido, particularmente, a que la máquina compuesta de Woolf, en la que se empleaba aquél, había sido introducida allí pocos años antes. Así pues, la pretensión expuesta por Carnot en las *Réflexions* de deducir la máxima potencia de una máquina de vapor resultaba muy oportuna.

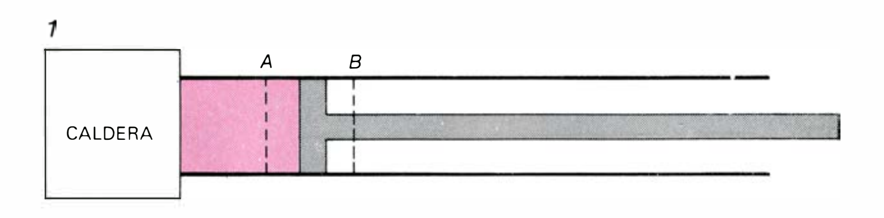
Las *Réflexions* se proponían dar pública noticia de las posibilidades de la máquina de vapor para mejorar el nivel de vida en Francia. La obra empieza con un detallado informe acerca de la importancia de la máquina de vapor, sobre todo en Inglaterra. Carnot escribe:

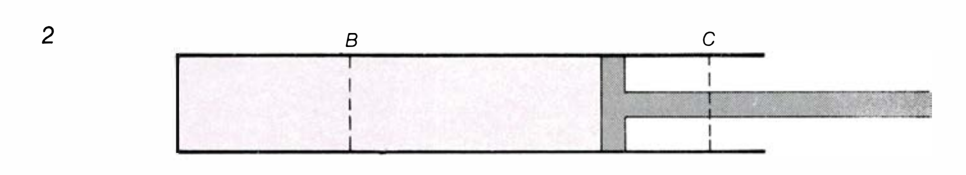
"El estudio de estas máquinas es del máximo interés, enorme su importancia y, su utilización, cada vez mayor. Parecen destinadas a producir una gran revolución en el mundo civilizado.

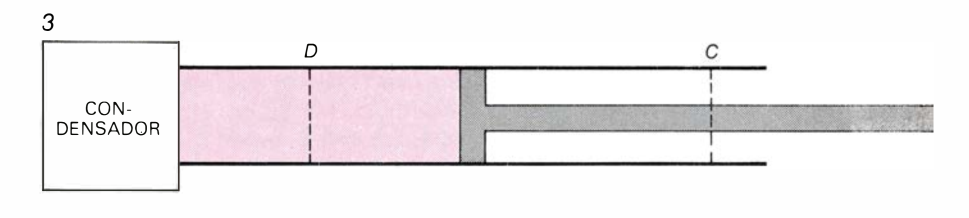
"La máquina de vapor explota ya nuestras minas, mueve nuestras naves, excava nuestros puertos y ríos, forja el hierro, da forma a la madera, muele el grano, hila y teje nuestros tejidos, transporta las más pesadas cargas, etcétera. Parece que algún día deberá servir de motor universal y ser preferida a la fuerza de los animales, los saltos de agua y corrientes de aire. Tiene, sobre el primero de estos motores, la ventaja de la economía; sobre los otros dos, el inapreciable valor de poder emplearse en todas las épocas y lugares y de no sufrir ninguna interrupción en su trabajo.

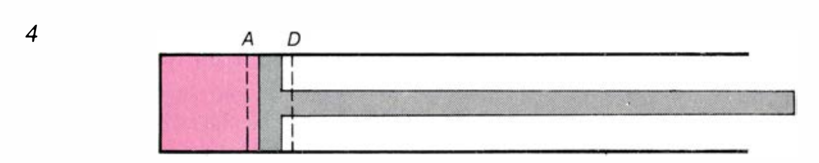
"Si algún día los perfeccionamientos de la máquina de vapor la hacen poco costosa de instalación y combustible, reunirá todas las cualidades deseables y promoverá un desarrollo de las artes industriales cuyo alcance será difícil predecir. En efecto, no sólo se trata de un motor potente y cómodo, que puede obtenerse y transportar a todas partes, que sustituye a los motores ya en uso, sino de un motor que proporciona a las



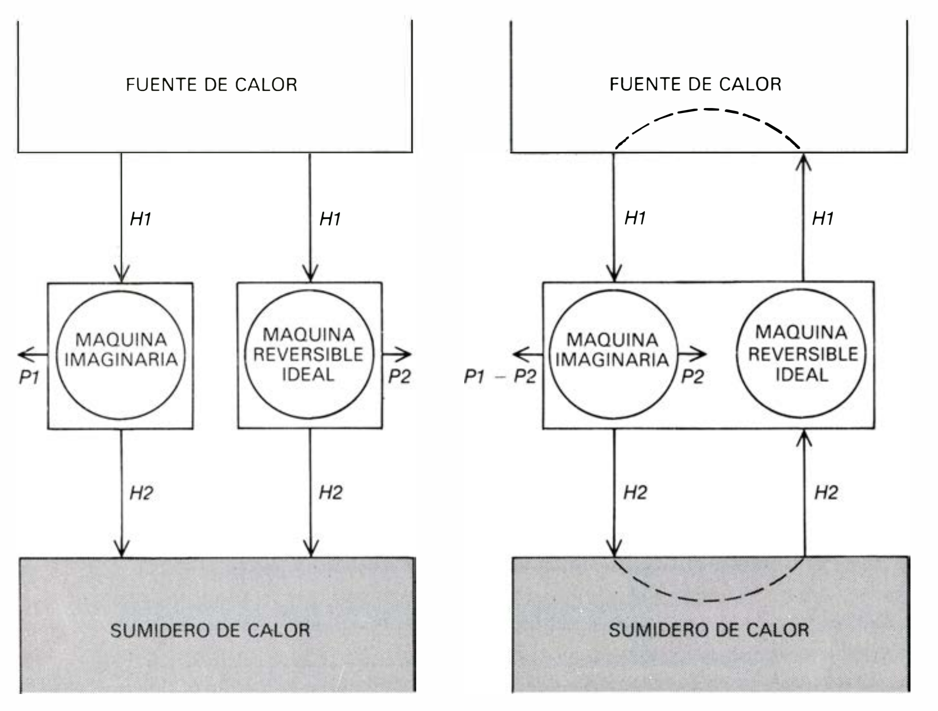








DE CUATRO FASES CONSTA EL CICLO DE CARNOT, representada cada una de ellas por una línea en la gráfica presión-volumen y por un diagrama de un pistón en un cilindro. Las letras de la gráfica corresponden a las posiciones marcadas con letras en el pistón. En la primera fase, se pone al cilindro en contacto con una fuente de calor, como una caldera. Cuando el vapor del cilindro se expande, absorbe bastante calor de la caldera para mantener una temperatura y una presión constantes. En la segunda fase se quita la caldera. Como el vapor continúa su expansión, bajan la presión y la temperatura. En la tercera fase se pone el cilindro en contacto con un cuerpo a baja temperatura, un condensador, por ejemplo. Cuando se condensa el vapor, da bastante calor al condensador para mantener constantes la presión y la temperatura. En la cuarta fase se quita el condensador. Puesto que se continúa comprimiendo el vapor, aumentan su presión y su temperatura. Al alcanzar ésta su valor inicial, el ciclo se completa y la máquina torna a su estado original. Ningún ciclo genera más potencia motriz que el de Carnot. Sadi no demostró, empero, que el ciclo tuviera el rendimiento térmico máximo posible: la razón máxima del trabajo realizado al calor consumido. La primera ley de la termodinámica, que expresa la equivalencia entre el trabajo y el calor como formas de energía, no se había establecido. La evaluación del rendimiento térmico del ciclo de una máquina térmica requiere la noción de temperatura absoluta, desconocida entonces también.



EN LA REVERSIBILIDAD TERMODINAMICA se apoyó Carnot para probar que ningún ciclo podía producir un efecto útil mayor que su ciclo. En las *Réflexions*, consideró al calor como un fluido indestructible conocido como calórico. Ordinariamente, el calórico fluye de un cuerpo caliente a otro más frío; si se pudiera invertir el ciclo, el calórico podría ser transportado de un cuerpo con la temperatura baja a otro con la temperatura alta y el ciclo consumiría tanta potencia motriz cuanta genera cuando funciona directamente. La prueba se inicia suponiendo que hay un ciclo que podría producir más potencia motriz. A la izquierda, está el diagrama del ciclo imaginario y del ciclo reversible ideal operando entre una fuente de calor y un sumidero de calor. *P1* es la potencia motriz del ciclo imaginario y *P2* es la del ciclo reversible; *H1* y *H2* representan el flujo de calor. A la derecha figura un diagrama del ciclo imaginario conduciendo al ciclo ideal en sentido inverso. Por ser *P1* mayor que *P2*, hay una potencia residual para efectuar trabajo exterior y el sistema constituye una máquina de movimiento perpetuo. La posibilidad de movimiento perpetuo había sido rechazada desde hacía tiempo (por el padre de Sadi, entre otros), y, así, Carnot concluyó que el supuesto de un ciclo más potente que el ideal reversible era incorrecto, por absurdo.

artes a las que se aplica una extensión rápida y que puede incluso crear artes enteramente nuevas.

"El servicio más notable que la máquina de vapor ha rendido en Inglaterra es, sin duda, el de haber reanimado la explotación de sus minas de hulla, que había disminuido y que amenazaba con extinguirse enteramente a causa de la creciente dificultad de los drenajes y de la extracción del combustible. En segundo lugar, hay que situar los servicios rendidos en la fabricación del hierro, tanto por la hulla, obtenida en abundancia, que ha sustituido a la madera en el momento en que ésta comenzaba a escasear, como por las potentes máquinas de todas clases cuyo uso ha posibilitado o facilitado la máquina de vapor.

"El hierro y el calor son, como se sabe, los alimentos, los soportes de las artes mecánicas. Posiblemente no existe en Inglaterra ninguna industria cuya existencia se base en el uso de estos agentes y que no los emplee con profusión. Quitarle hoy a Inglaterra sus máquinas de vapor, sería dejarla a la vez sin hulla y sin hierro; sería secar todas sus fuentes de riqueza, arruinar todos sus medios de prosperidad; sería aniquilar esta colosal potencia. La destrucción de su marina, que ella considera su baluarte más firme, le resultaría probablemente menos funesta.

"La navegación segura y rápida de los barcos de vapor puede considerarse un arte enteramente nuevo, debido a este tipo de máquina. Este arte ha hecho posible ya el establecimiento de comunicaciones a través de los brazos de mar y sobre los grandes ríos del antiguo y del nuevo continente. Ha permitido recorrer regiones aún salvajes, donde hasta hace poco tiempo difícilmente se podía penetrar; ha hecho posible llevar los frutos de la civilización a puntos del globo donde los habrían tenido que esperar muchos años, si no hubiera sido por ese ingenio. La navegación con máquinas de vapor de algún modo acerca entre sí a las naciones más lejanas. Tiende a reunir a los pueblos de la tierra como si habitasen un mismo país. En efecto, disminuir el tiempo, las fatigas, las incertidumbres y los peligros de los viajes, ¿no equivale a abreviar mucho las distancias?"

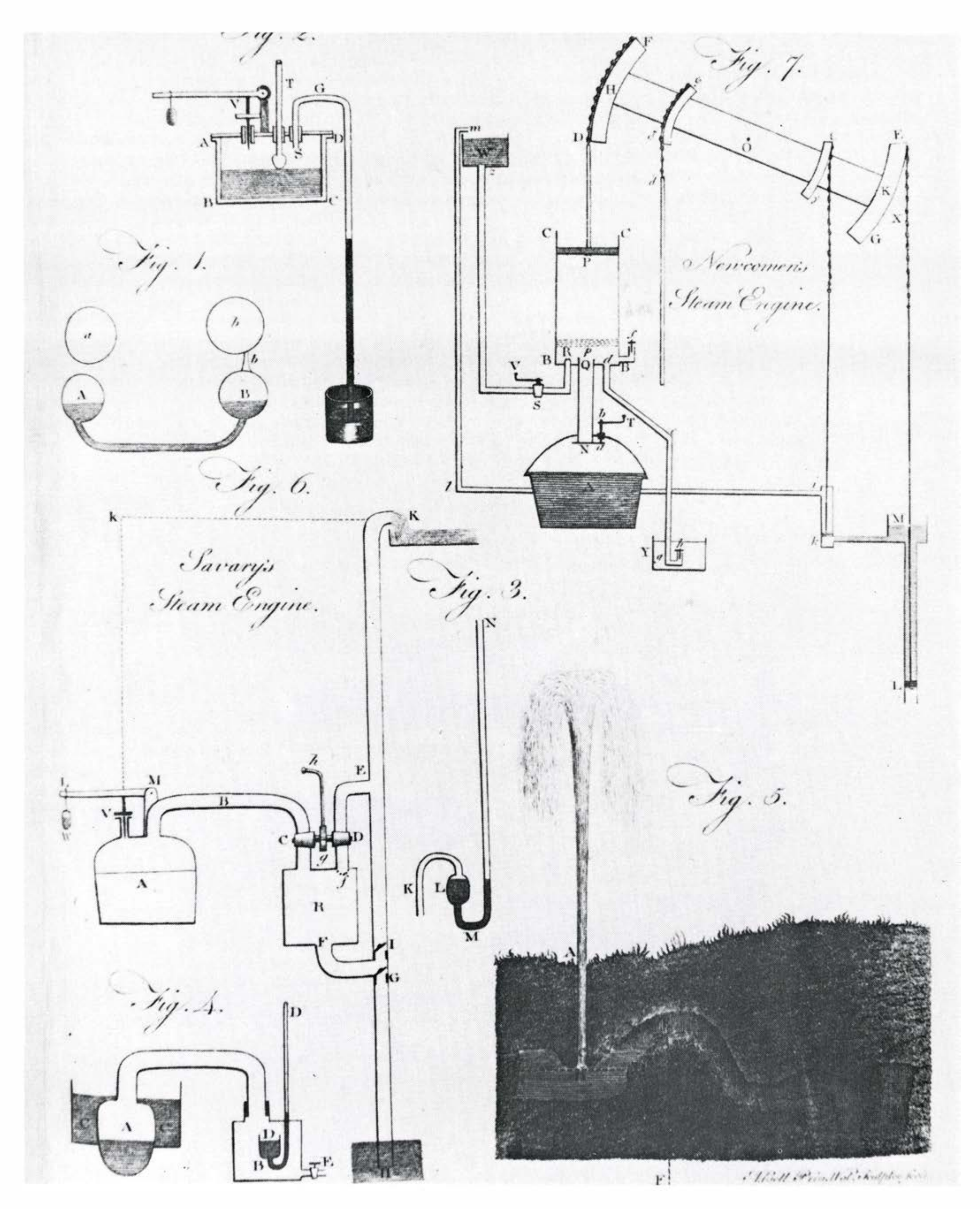
Carnot procede entonces a exponer el desarrollo empírico e histórico de la máquina de vapor, mostrando que estaba bien informado de la contribución debida a cada ingeniero inglés. Curiosamente, ignoró el trabajo de su compatriota Denis Papin, que inventó el autoclave de vapor u olla a presión y propuso una bomba de vapor. El relato continúa:

"Como la mayoría de los hallazgos humanos, el descubrimiento de las máquinas de vapor tiene su origen en ensayos rudimentarios, ensayos atribuidos a diversas personas y cuyo verdadero autor se desconoce. Por lo demás, el principal descubrimiento no corresponde a estos ensayos, sino a los perfeccionamientos sucesivos que han conducido a las máquinas de vapor hasta el estado en que las encontramos hoy en día. Hay, poco más o menos, tanta distancia entre los primeros aparatos en los que se desarrolló la fuerza expansiva del vapor y las máquinas actuales como entre la primera balsa que los hombres construyeron y las modernas embarcaciones.

"Si el honor de un descubrimiento corresponde a la nación en la que éste se acunó y creció, ese honor no se le puede negar a Inglaterra: Savery, Newcomen, Smeaton, el célebre Watt, Woolf, Trevithick y algunos otros ingenieros ingleses son los verdaderos creadores de la máquina de vapor. En sus manos ha adquirido todos los sucesivos grados de perfeccionamiento. Por lo demás, es natural que una invención nazca y, sobre todo, se perfeccione allí donde su necesidad se hace sentir más imperiosamente."

Los acontecimientos subsiguientes en el desarrollo de la máquina de vapor y otras máquinas térmicas confirmaron la última observación de Carnot en el texto citado. Mientras escribía las *Réflexions*, algunos ingenieros de Cornwall, de los que el más conocido era Richard Trevithick, introdujeron notables mejoras en las máquinas de amplio uso en la mencionada localidad para el bombeo de agua de las minas. Thomas Newcomen, Watt y otros investigadores tenían en la mente esta aplicación cuando diseñaron máquinas de bombeo de vapor.

Las mejoras que condujeron a una máquina particularmente fructífera, la máquina de Cornish, se derivaron de varios factores; entre ellos: la necesidad de ahorrar carbón, que tenía que transportarse a Cornwall, la expiración, en 1800, de una patente que Watt



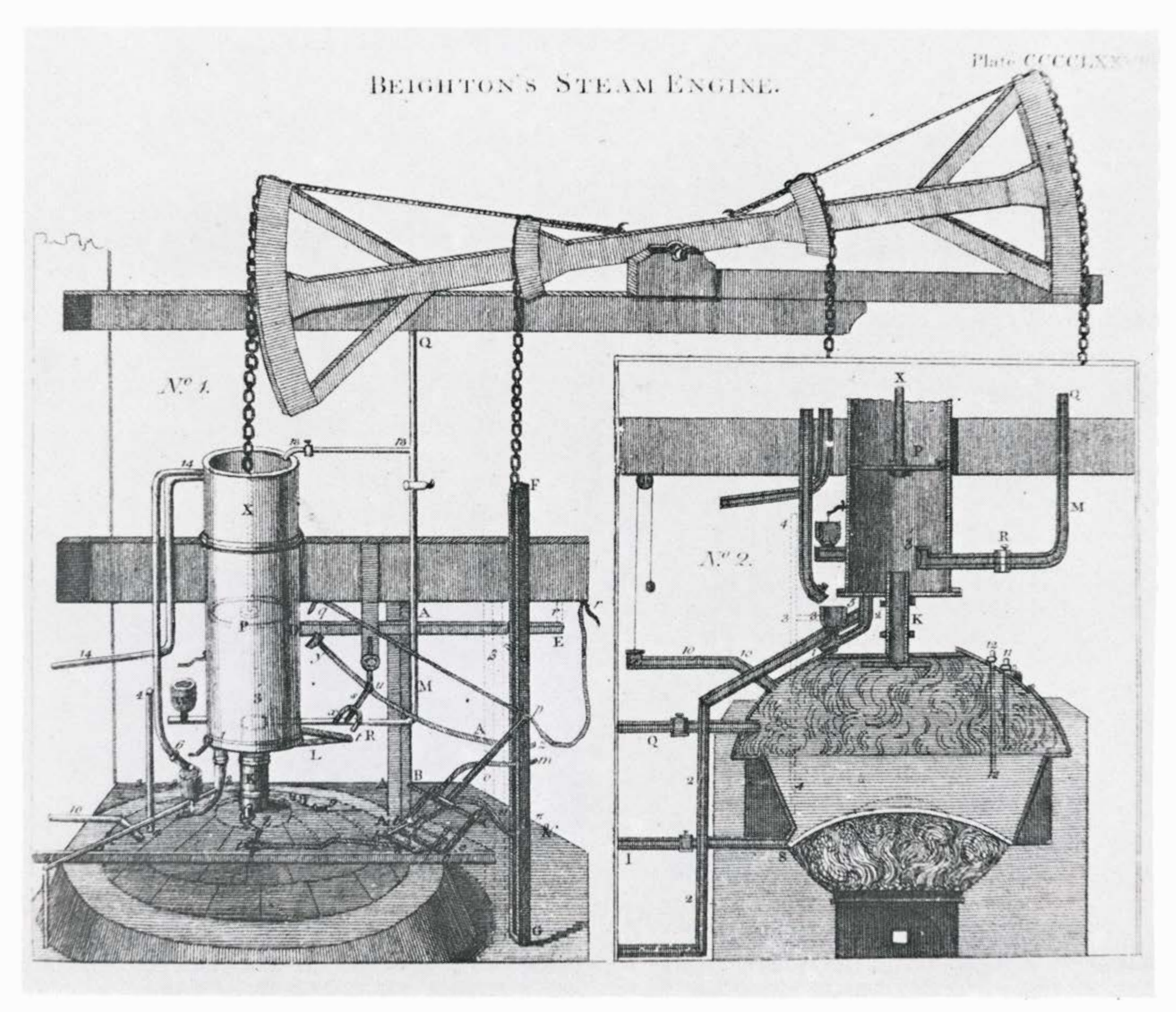
MAQUINAS DE BOMBEO por medio del vapor de agua diseñadas por Thomas Savery y Thomas Newcomen. Este esquema de las mismas apareció reproducido en 1797 en la tercera edición de la *Encyclopaedia Britannica*. En 1698, Savery había obtenido una patente por una "nueva Invención para Elevar el Agua y producir Movimiento en cualquier clase de máquina por la Fuerza Impelente del Fuego, que será de gran utilidad y Conveniencia para el Drenaje de Minas, Abastecimiento de Agua de las Ciudades y para el Trabajo de toda Clase de Molinos que carecen del Beneficio del Agua y de Vientos constantes". El diagrama muestra una versión mejorada de la máquina de Savery, que bombea el agua por medio de la succión producida por la condensación del vapor en una vasija cerrada. En 1712, Newcomen desarrolló la primera máquina de vapor a pistón. El vapor de agua se producía en la vasija situada debajo del cilindro y entraba en el cilindro a través de la válvula *N*. Para condensar el vapor, se inyectaba agua fría en el cilindro a través de la válvula *S*. En las *Réflexions* Carnot cita la *Encyclopaedia Britannica* como una fuente de información sobre varias clases de máquinas movidas por vapor de agua.

y Mattew Boulton habían poseído sobre un método de ahorrar combustible condensando vapor, y la publicación, iniciada en 1810, de *Engine Reporter* por Joel Lean y sus parientes. *Engine Reporter* incluía estadísticas sobre cuánta agua bombeaban las mejores máquinas de Cornish por "bushel" de carbón (un bushel de carbón equivale a 42,582 kilogramos). Las cantidades, en millones de libra-pies (una libra-pie equivale aproximadamente a 0,14 kilográmetros) de agua bombeada, pasaron de 15,7 en 1811 a 20,5 en 1814 y a 28 en 1820. La mejor máquina alcanzó 100 millones de libra-pies para 1835 y una máquina dio un promedio de 107 para un mes de 1842.

En las *Réflexions*, Carnot se ocupó seguidamente del análisis general de la potencia de las máquinas de vapor y otras máquinas térmicas. Trató de determinar si había un límite para la potencia motriz del calor, que impediría el progreso indefinido de las máquinas térmicas. También se preguntó si alguna substancia termodinámica, como el alcohol o el aire, podría ser preferible al vapor. Estas cuestiones ya habían sido examinadas por otros investigadores, pero sólo con respecto a un diseño particular. Carnot intentó hallar respuestas que pudieran aplicarse a todas las máquinas térmicas imaginables.

Carnot fue capaz de plantear estas cuestiones, a pesar de que las leyes de la termodinámica le eran desconocidas. La primera ley, que establece que en un sistema termodinámico la energía se conserva, no fue formulada hasta la década de 1840 por James Prescott Joule. Aspecto fundamental de la misma es la equivalencia entre el calor y el trabajo como formas de energía. Esta equivalencia se oponía a la teoría del calórico, que, a pesar de alguna oposición, seguía predominando cuando Carnot escribió las *Réflexions*.

Según la teoría del calórico, el calor es un fluido imponderable y sin masa, que se conserva siempre. En favor del genio de Carnot abona el hecho de que pudiera extraer conclusiones correctas y de gran transcendencia sobre las máquinas térmicas sin suponer que el calor podía convertirse en trabajo mecánico y el trabajo mecánico en calor. Sin conocer la primera ley de la termodinámica no podía definir el rendimiento térmico de una máquina, como se afirma gratuitamente que hizo. Carnot sólo definió el máximo "servicio" de una



MAQUINA ATMOSFERICA investigada por Henry Beighton, según una ilustración procedente de la *Britannica* de 1797 proporcionada por S. L. Rabinowitz. Beighton hizo el primer análisis cuantitativo de la potencia motriz de una máquina de vapor. Calculó las profundidades desde las que máquinas de cilindros de diversos tamaños podrían extraer agua de una mina. Los cálculos de Beighton fueron publicados en el *The Ladies'Diary*, del que era editor.

máquina: su efecto útil usual (cuánta agua es elevada multiplicado por la altura a la que es elevada) para una cantidad de carbón dada.

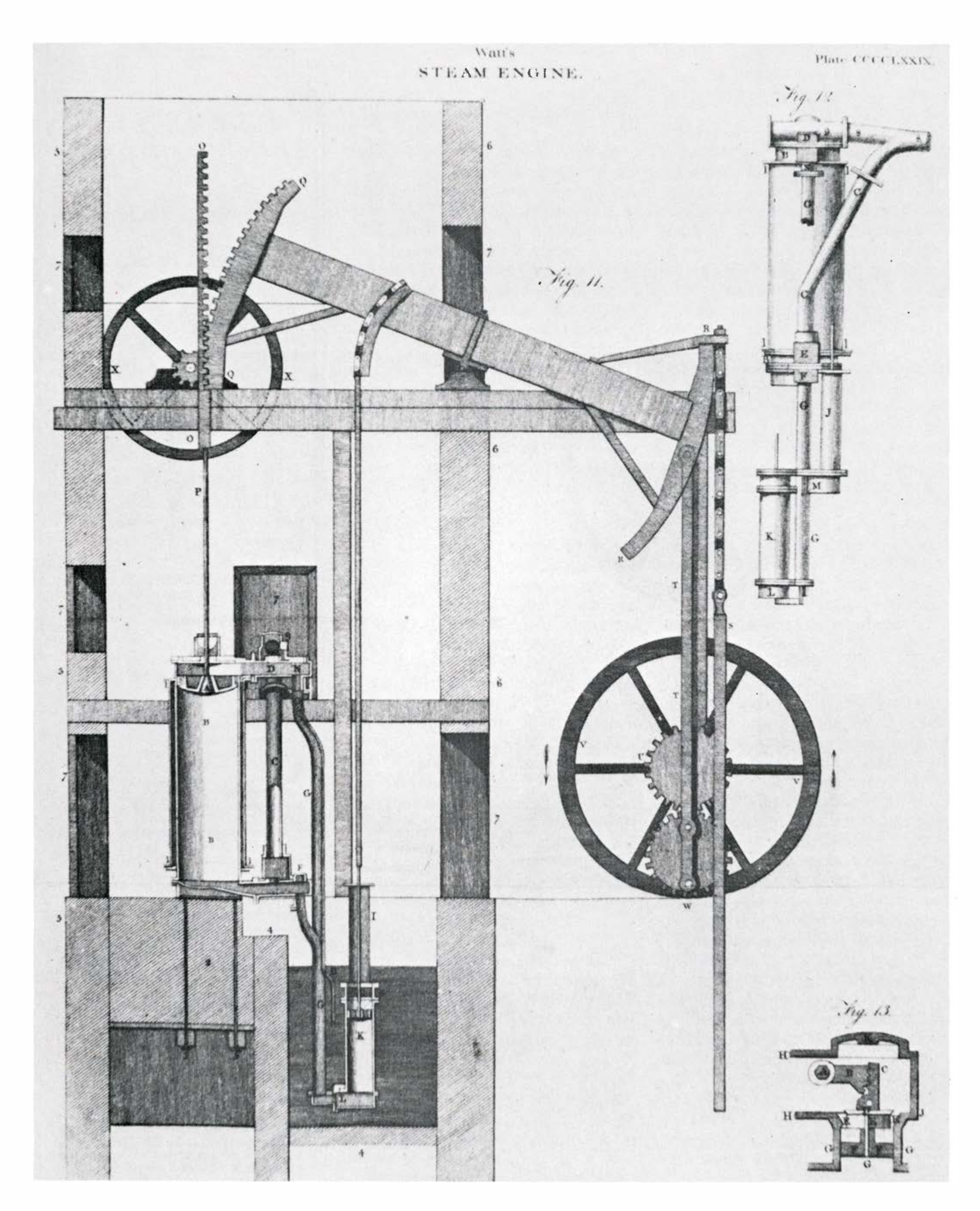
Carnot supuso que la potencia motriz de una máquina térmica procede del paso de calórico desde un cuerpo caliente (un depósito de calor o caldera) hasta otro frío (un sumidero de calor o condensador). Lazare había publicado un tratado en el que analizaba el rendimiento de las máquinas movidas por agua y, en las *Réflexions*, Sadi estableció una analogía entre la máquina de vapor y la rueda hidráulica. El trabajo realizado por una máquina térmica resultaba del paso del calórico a través de una diferencia de temperaturas, del mismo modo como el trabajo realizado por una rueda hidráulica resultaba del paso del agua a través de una diferencia de potencial gravitatorio. En la rueda hidráulica, la potencia motriz dependía de la cantidad de agua y de la distancia desde la que el agua cae. En la máquina térmica, la potencia motriz dependía de la cantidad de calórico y de la magnitud de la diferencia de temperaturas entre las que se mueve el calórico.

Carnot describió un ciclo ideal para una máquina térmica compuesta de un cilindro, un pistón, una substancia termodinámica, una fuente de calor y un sumidero de calor. La substancia, que podía ser cualquier material (sólido, líquido o gas) y que podía someterse a expansiones y contracciones mediante aplicaciones sucesivas de calor y frío, se situaba en el cilindro, dentro del cual se había ajustado convenientemente el pistón. El ciclo tenía cuatro fases. En la primera, se ponía la substancia en contacto con la fuente de calor, de suerte que, cuando dicha substancia se expandía, tomaba bastante calor del depósito para mantener su temperatura constante. En la segunda, se aislaba del depósito de calor la substancia, con lo que, al continuar ésta su expansión, su temperatura disminuía. En la tercera, se ponía la substancia en contacto con el sumidero de calor, de modo que, al comprimirse, liberaba suficiente calor para mantener constante su temperatura. En la cuarta, se aislaba la substancia del sumidero, con lo que, al proseguir su compresión, aumentaba su temperatura. Cuando la temperatura tornaba a su valor inicial, se cerraba el ciclo y la substancia y la máquina habrían recobrado su estado original. El ciclo podía repetirse indefinidamente para producir potencia sin cesar. Expresado en términos actuales, el ciclo constaba de una expansión isotérmica (desarrollada a temperatura constante), una expansión adiabática (en la que el calor se convertía en trabajo), una compresión isotérmica y una compresión adiabática.

Al introducir el ciclo ideal, Carnot escribe: "Imaginemos dos cuerpos *A* y *B* mantenidos ambos a una temperatura constante, siendo la de *A* más elevada que la de *B*: estos dos cuerpos, a los que se les puede dar o quitar calor sin hacer variar su temperatura, desempeñarán las funciones de dos depósitos indefinidos de calórico. Llamaremos al primero hogar y al segundo refrigerante."

En una máquina térmica actual, la fuente de calor es casi siempre finita; así, cuando se consume el calor la temperatura de la fuente disminuye. Por este motivo, la mayoría de las máquinas térmicas prácticas tienen un proceso de admisión del calor que no es isotérmico. Esto vale para todas las máquinas de combustión interna y turbinas de gas, e incluso en la máquina de vapor su desarrollo se ha ido apartando del proceso de admisión del calor casi completamente isotérmico de las primeras máquinas. En su lugar, los diseñadores modernos prefieren el ciclo de vapor a alta presión, en el que la parte isotérmica de la admisión del calor es pequeña, y el ciclo de vapor supercrítico, que no tiene admisión de calor isotérmica. La idea de un sumidero de calor infinito tiene más visos de verosimilitud que la de una fuente de calor infinita: suele haber una ilimitada provisión de aire o agua para la refrigeración. Por esta razón, un ciclo basado en una expulsión de calor isotérmica es accesible, y se lleva a cabo en la mayoría de los ciclos de vapor que condensan el vapor a temperatura constante.

Con su ciclo, Carnot introdujo una noción termodinámica fundamental, la de reversibilidad. Cada repetición del ciclo implicaba el flujo de calórico de un cuerpo con temperatura elevada a un cuerpo con temperatura baja, con la consiguiente producción de trabajo. Como se pensaba que el calórico era un fluido indestructible, Carnot razonó que cada fase del ciclo podía invertirse aportando trabajo a la máquina, de modo que el calórico sería transportado desde el cuerpo de baja temperatura hasta el cuerpo de temperatura alta. En la rueda hidráulica, usada como analogía, el proceso podía invertirse aportando trabajo al agua, de modo que ésta se elevaba, a través de la diferencia de potencial gravitatorio, hasta



MAQUINA DE WATT, desarrollada en 1765 por James Watt. Era en esencia una máquina Newcomen en la que el vapor se condensaba, no en el cilindro, sino en un recipiente aparte. La introducción de un condensador separado fue de la mayor importancia debido a que el calor ya no se consumía en el calentamiento y enfriamiento alternativo del cilindro, manteniéndose el cilindro y el pistón a la temperatura del vapor. El condensador separado produjo un ahorro de combustible de casi el 75 por ciento. A principios de los años 1780, Watt hizo la primera máquina que operaba tanto en el movimiento ascendente como en el descendente del pistón. También hizo una de las primeras máquinas que convertía movimiento alternativo en movimiento de rotación. En la máquina de la ilustración, la conversión se realiza por un sistema de engranajes en el que un engranaje planeta (*W*) en el extremo de una rueda (*T*), describe una órbita alrededor de un engranaje sol (*U*). El engranaje sol, cuyo movimiento lo facilita un volante (*V*), proporciona movimiento de rotación.

su fuente, en la parte superior de la rueda. Puesto que cada fase del ciclo de Carnot podía invertirse, el ciclo entero resultaba reversible. Al invertir su curso, consumía tanta potencia motriz cuanta generaba en el proceso directo.

El eje de la argumentación teórica de las *Réflexions* es la demostración de Carnot de que, para una diferencia de temperatura dada y una determinada cantidad de calórico, ningún ciclo podía generar más potencia motriz que su ciclo ideal o reversible. Imagínese un ciclo que pudiera generar más potencia. Este ciclo podría disponerse para conducir el ciclo ideal en sentido inverso y, para volver a su estado original el ciclo ideal, sólo necesitaría consumir una parte de la potencia motriz producida por el ciclo imaginario. El resto de la potencia quedaría disponible para realizar trabajo exterior. Expresándolo de otro modo, si se pudiera conectar el ciclo reversible ideal a un ciclo con más rendimiento, tendríamos una máquina de movimiento perpetuo.

La posibilidad de una máquina de movimiento perpetuo había sido rechazada desde tiempo atrás por la mayoría de los investigadores, incluido Lazare Carnot, quien había escrito sobre el tema. En consecuencia, Sadi Carnot sacó la conclusión de que el supuesto inicial de un ciclo que pudiera generar más potencia motriz que el ciclo ideal o reversible era incorrecto. Con esta demostración Carnot expresó en esencia la segunda ley de la termodinámica, que no recibiría formulación explícita hasta los trabajos de Clausius y Thomson en la década de 1850. Según esta ley, no es posible ningún proceso cuyo único resultado sea la absorción de calor del depósito y la conversión de calor en trabajo. La ley indica que un proceso termodinámico espontáneo sólo puede desarrollarse en una dirección, desde una fuente de temperatura alta hasta un sumidero o refrigerante de temperatura baja. Para invertir el proceso se requiere energía procedente de una fuente exterior.

Carnot demostró que la potencia del ciclo ideal no dependía de la substancia termodinámica utilizada, aunque la naturaleza de ésta era importante desde un punto de vista práctico, en la medida en que sus propiedades afectaban al diseño de la máquina y a la temperatura a la que podía operar. Concluyó que la mejor substancia termodinámica era la que más se expandía para un cambio de temperatura dado. Con este criterio, los gases ofrecían mejores expectativas que los sólidos o los líquidos.

Carnot reconoció la ventaja potencial del aire como substancia termodinámica. Con el aire existía la posibilidad de combustión interna, lo que hacía innecesaria la caldera o cualquier otra forma de fuente exterior de calor. Tampoco se precisaba condensador. Carnot sabía (y aprobó el hecho) que un proceso de descarga como éste se estaba explotando ya, no con máquinas de aire, sino con máquinas de vapor a alta presión, siendo Trevithick el pionero de tales ingenios. La ausencia de condensador posibilitó el desarrollo de una máquina compacta y de fácil transporte y, en virtud de ello, de la locomotora de vapor. Carnot comprendió que el mecanismo que comprimía el aire en una máquina de aire debería ser más elaborado y consumir más potencia que el mecanismo correspondiente de una máquina de vapor. El agua era virtualmente incompresible y ocupaba un volumen pequeño si se comparaba con el volumen de la masa de vapor equivalente; por lo tanto, se necesitaba poca potencia para mover la bomba que alimentaba de agua la caldera.

Las ideas de Carnot acerca del aire como substancia termodinámica prefiguraron el desarrollo del motor de combustión interna: "La utilización del aire atmosférico para el desarrollo de la potencia motriz del calor", escribió, "presentaría en la práctica grandes dificultades, aunque quizá no insuperables. Si se llegase a vencerlas, el aire ofrecería notables ventajas sobre el vapor de agua." Desde los años 1790 se habían ensayado gases y pólvoras como combustibles, pero no fue hasta 1859 cuando una máquina de combustión interna, realizada por Jean Joseph Étienne Lenoir, funcionó con facilidad, y ello, a su vez, 17 años antes de que Nikolaus August Otto desarrollara el primer motor de combustión interna con éxito comercial. Las predicciones de Carnot, tanto de las dificultades como de las ventajas potenciales de la combustión interna, se confirmaron.

Carnot previó también la posibilidad de quemar el combustible de un motor de combustión interna por compresión. En 1891, Rudolf Diesel concibió un motor que funcionaba según un ciclo de Carnot con compresión por ignición. Diesel esperaba lograr la combustión interna controlando la razón a la que se inyectaba el combustible. No obstante, el primer motor que construyó en 1898 operaba según un ciclo diferente, como todos los subsiguientes motores diesel, en el que no había ningún proceso isotérmico. La historia del motor diesel confirma la idea de que el ciclo de Carnot no es la base de un motor actual, sino un ciclo hipotético con el que pueden estudiarse nociones termodinámicas tales como la máxima potencia motriz de una máquina.

Carnot también bosquejó el caso de un ciclo compuesto. "Podríamos incluso concebir la posibilidad de hacer actuar al mismo calor sucesivamente sobre el aire y sobre el vapor de agua. Bastaría dejar el aire, después de su uso, a una temperatura elevada y, en lugar de expulsarlo inmediatamente a la atmósfera, hacer que rodeara una caldera de vapor, como si acabase de salir directamente de un hogar." Se aprovechó un ciclo compuesto así, en la década de 1920, en la máquina Still, un motor diesel que incorporaba una caldera de escape y una máquina de vapor alternativa. La máquina Still, aunque probada en barcos y locomotoras, no llegó a popularizarse. Más recientemente se han introducido ciclos compuestos que combinan una turbina de gas con una turbina movida mediante vapor de agua u otro vapor. Tales ciclos compuestos se han ido estableciendo gradualmente, siglo y medio después de que Carnot los propusiera.

Carnot consideró otros vapores distintos del vapor de agua, pero los descartó debido a las dificultades prácticas y a que no ofrecían ventajas termodinámicas teóricas. Muy pocas máquinas han empleado con éxito otros vapores que no fueran el de agua y, ninguna, a gran escala. Hoy, empero, se está indagando en el movimiento de una máquina térmica con un vapor de alto peso molecular; en una turbina pequeña, una substancia termodinámica semejante ofrece ventajas prácticas sobre el vapor de agua.

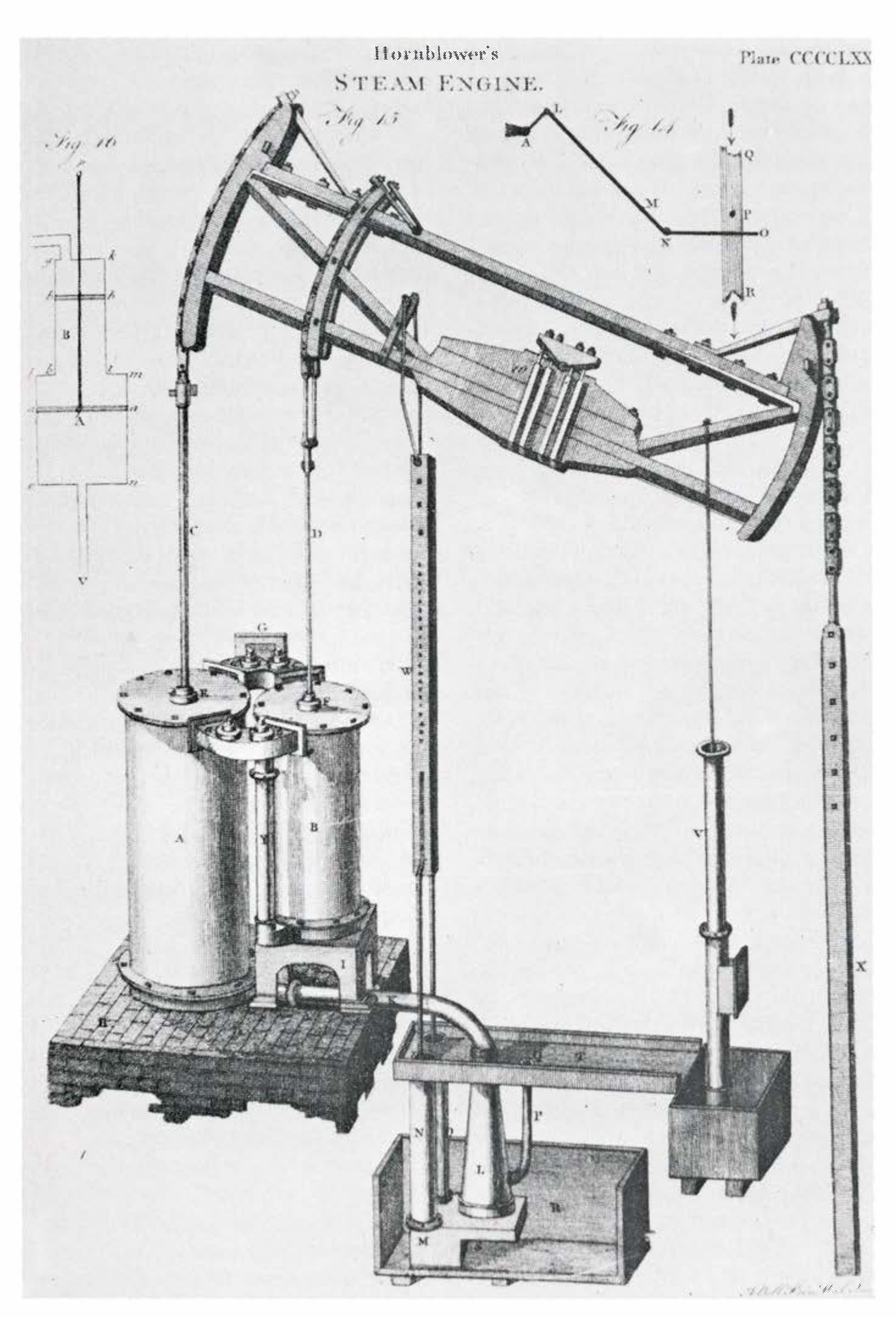
Aunque se le ha considerado fundador de la ciencia de la termodinámica, fue primero y por encima de todo un ingeniero, muy consciente de los compromisos que requería el diseño y la construcción de una máquina. Sus *Réflexions* terminan observando:

"El ahorro del combustible no es más que una de las condiciones a cumplir por las máquinas de vapor; en muchas circunstancias esta condición es secundaria y debe subordinarse a la seguridad, solidez, duración de la máquina, pequeño espacio a ocupar, pocos gastos de instalación, etcétera. Saber apreciar, en cada caso, en su justo valor, las consideraciones de conveniencia y economía que puedan presentarse; saber

discernir entre las más importantes y las que sólo son accesorias; saber ponderarlas todas entre sí debidamente para llegar por los medios más fáciles al mejor resultado: tal ha de ser la principal cualidad del hombre llamado a dirigir y a coordinar entre sí los trabajos de sus semejantes y hacerlos cooperar hacia un fin útil, cualquiera que sea."

Poco se sabe acerca de lo que hizo Carnot después de escribir las *Réflexions*. En 1828, trabajó durante un corto tiempo como ingeniero militar en Lyons y Auxonne, regresando luego a París para reemprender sus estudios sobre las máquinas y la teoría del calor. Una historia de la École Polytechnique, publicada en la época, nos lo describe como constructor de máquinas de vapor, pero no existen pruebas de que las construyera en realidad.

Carnot aprobó la revolución de julio de 1830, que forzó la abdicación de Carlos X, porque parecía prometer un régimen más liberal. De hecho se sugirió que Carnot sería invitado a entrar en el nuevo gobierno como miembro de la Chambre des Pairs. Sin embargo, la sugerencia no cristalizó, porque Carnot pronto se desilusionó con un gobierno que restauró la monarquía, rechazando la naturaleza hereditaria del cargo que se le había propuesto.



MAQUINA DE DOBLE EXPANSION, desarrollada en 1781 por Jonathan Hornblower. Constaba de dos cilindros. Después de que el vapor sufría la expansión en un cilindro pasaba a otro cilindro en el que podía continuar expandiéndose. En principio, la doble expansión generaría más potencia motriz que la expansión simple. En la práctica, los dos cilindros resultaron ser una complicación innecesaria porque la presión inicial del vapor de agua era demasiado baja. Ulteriormente se desarrollaron máquinas que empleaban efectivamente una expansión compuesta. La ilustración procede de la *Britannica* de 1797.

Poco después de estos acontecimientos, Carnot empezó a analizar la relación entre la temperatura y la presión de un gas, pero en junio de 1835 contrajo la escarlatina y la "fiebre cerebral". El 24 de agosto, a la edad de 36 años, murió a causa de una epidemia de cólera. Como era común en las víctimas del cólera, fue quemado con casi todos sus papeles y otras pertenencias.

Se salvaron algunos fragmentos de sus cuadernos de trabajo. En 1878, Hippolyte se dirigió de nuevo a la Academia de Ciencias francesa con el manuscrito de las *Réflexions*. Un fragmento, escrito al parecer entre 1824 y 1826, indica que Sadi había llegado a rechazar la teoría del calórico y a comprender la equivalencia entre calor y trabajo. "El calor no es otra cosa que la potencia motriz o, mejor, el movimiento", escribió. "Es un movimiento de las partículas de los cuerpos. Dondequiera que hay destrucción de la potencia motriz en las partículas de los cuerpos, hay al mismo tiempo producción de calor en cantidad exactamente proporcional a la cantidad de la potencia motriz destruida; recíprocamente, dondequiera que hay destrucción de calor, hay producción de potencia motriz." Llegó incluso a bosquejar experimentos que podrían demostrar la equivalencia entre calor y trabajo, pero nunca los llevó a cabo. Algunos de estos experimentos eran similares a los que Joule realizó, unos 15 años después, estableciendo la primera ley de la termodinámica.

Resulta interesante, aunque ocioso, especular sobre lo que Carnot podría haber realizado si hubiese vivido más tiempo. ¿Se habrían llevado a la práctica sus ideas sobre las máquinas térmicas? ¿O habría muerto como un profeta al que nadie presta oídos? No es demasiado tarde para rendir tributo a su perspicacia y a su capacidad analítica, firmemente asida a las realidades prácticas. La necesidad de atender a los intereses humanos en la vanguardia del trabajo científico y el reconocimiento de la naturaleza empírica del progreso tecnológico son lecciones que el investigador moderno debería tener presentes.

Juegos matemáticos

La parábola: curva abstracta que se adapta bien al mundo concreto

Martin Gardner

Los matemáticos de todos los tiempos se han dedicado a construir objetos abstractos y a examinar sus propiedades por la única razón de encontrarlos bellos e interesantes. Más tarde, siglos quizá, puede suceder que tales objetos resulten de enorme utilidad al aplicarlos al mundo físico. Ningún ejemplo refleja mejor este hecho que los trabajos realizados en la Grecia clásica sobre el tema de las cuatro secciones cónicas. Ya me he referido a tres de ellas en artículos anteriores: circunferencias, elipses e hipérbolas. Esta vez echaremos un vistazo a las parábolas.

Al interceptar un cono de revolución por un plano perpendicular a su eje, la sección transversal producida es una circunferencia. Basta darle al plano una leve oblicuidad con respecto al eje de rotación para que la sección se transforme en una elipse, que es el lugar geométrico de todos los puntos del plano cuyas distancias a dos puntos fijos, los focos, tienen suma constante. Podemos imaginar las circunferencias como elipses cuyos focos se hubieran fundido en un único punto, el centro de la circunferencia. Conforme el plano de corte va situándose más y más oblicuamente al eje del cono, los focos de la elipse van separándose progresivamente, y las elipses van acentuando su "excentricidad". En el momento en que el plano llegue a situarse paralelamente a una (y sólo una) de las generatrices del cono, la sección producida será una parábola. Las parábolas, como las circunferencias, son curvas límite de elipses; en ellas, ha desaparecido uno de los focos, trasladándose hasta el infinito. Es una elipse, que, con palabras de Henri Fabre, "busca vanamente su segundo centro, ya perdido sin remedio".

Al ir siguiendo los brazos o ramas de la parábola en su huida hacia el infinito, vemos que tienden a situarse paralelamente una a otra, sin jamás conseguirlo (salvo en el infinito). He aquí cómo lo expresó Johannes Kepler al examinar las secciones cónicas:

"A causa de su naturaleza intermedia, la parábola ocupa posición central (entre la elipse y la hipérbola). Conforme se va produciendo, no separa sus brazos como la hipérbola, sino que va juntándolos y situándolos cada vez más paralelos, abarcando siempre un poco más, pero ambicionando también cada vez menos; la hipérbola, en cambio, cuanto más abarca más trata de lograr."

Una parábola es el lugar geométrico de todos los puntos de un plano cuya distancia a una recta fija (directriz) es igual a su distancia a un punto fijo (foco) no situado sobre aquélla. La ilustración de esta página muestra la forma habitual de situar una parábola respecto de unos ejes rectangulares de coordenadas cartesianas a fin de lograr que su ecuación sea lo más sencilla posible. Observemos que el eje de la parábola pasa por el foco y corta perpendicularmente a la directriz y, también, que la extremidad de la curva —el vértice— se encuentra en el punto (0,0), que es el origen de coordenadas. La cuerda trazada perpendicularmente al eje que pase por el foco es la longitud focal, llamada también lado recto. Sea a la distancia del foco al vértice. Como es evidente, también la distancia del vértice a la directriz será a. No es difícil demostrar que el lado recto ha de medir entonces $4a$. La parábola queda ahora descrita como el lugar de puntos del plano cartesiano cuyas coordenadas cumplen la ecuación canónica $y^2 = 4ax$, que en los manuales españoles suele escribirse $y^2 = 2px$, donde $p = 2a$ es el parámetro, o sea, la distancia foco-directriz. En general, toda ecuación de segundo grado de la forma $x = Ay^2 + By + C$ tiene por representación gráfica una parábola, aunque su posición con relación a los ejes no será necesariamente la típica que vemos en la ilustración.

Una sorprendente propiedad de las parábolas es la de que todas tengan igual forma. Desde luego, arcos de parábola como los mostrados en la ilustración superior de la página siguiente tienen formas diferentes; sin embargo, imaginando uno cualquiera de los arcos prolongado hasta el infinito, podemos tomar el otro, efectuar en él un cambio de escala conveniente y adosarlo después sobre algún lugar

Parábola cuyo vértice se encuentra en el origen de coordenadas y cumple la ecuación canónica $y^2 = 4ax$

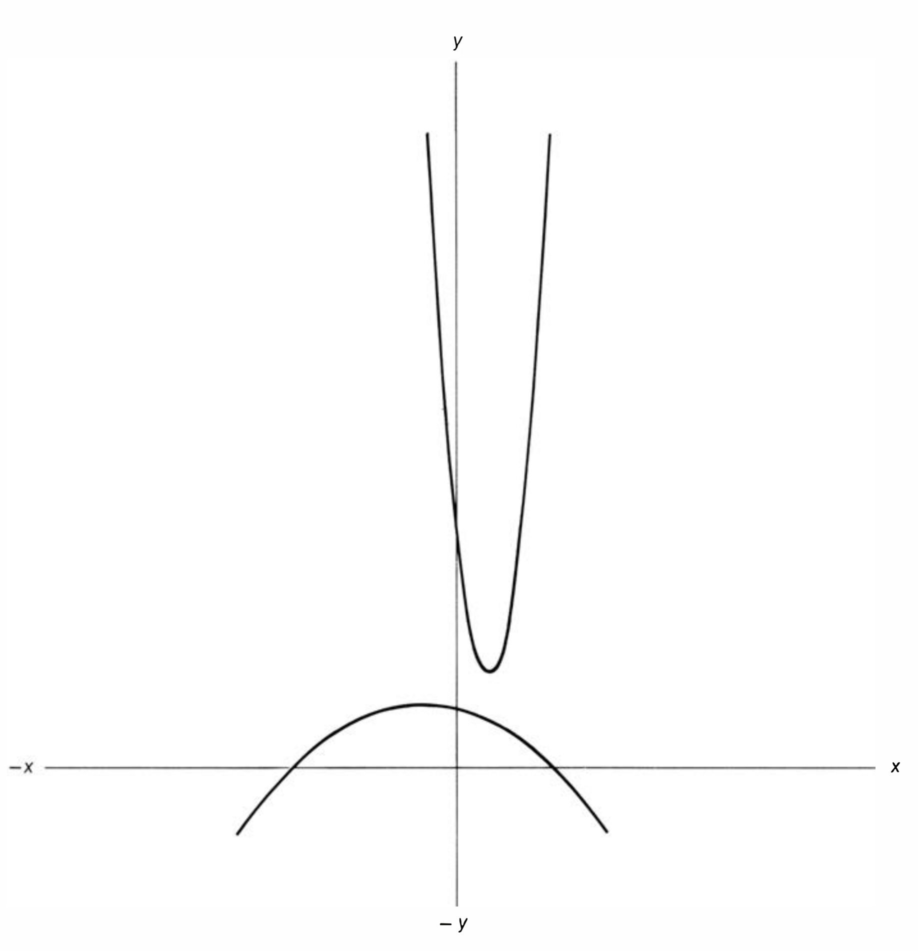
idóneo del primero, donde encajará perfectamente.

Esta propiedad de variar únicamente de tamaño, pero no de forma, es una de las que poseen en común circunferencias y parábolas, y que no poseen en cambio elipses ni hipérbolas. Todas las circunferencias son semejantes, porque, individualmente, todos los puntos son semejantes. Todas las parábolas son semejantes, porque todas las figuras formadas por una recta más un punto no perteneciente a ella son semejantes. Dicho de otro modo, cualquier par foco-directriz puede ser hecho coincidir con cualquier otro tras ejecutar una homotecia, una rotación y una traslación adecuadas. Cualquier parábola trazada sobre papel de gráficos debidamente graduado, y convenientemente posicionada en él, puede admitir por ecuación cualquier fórmula de segundo grado del tipo $x = Ay^2 + By + C$.

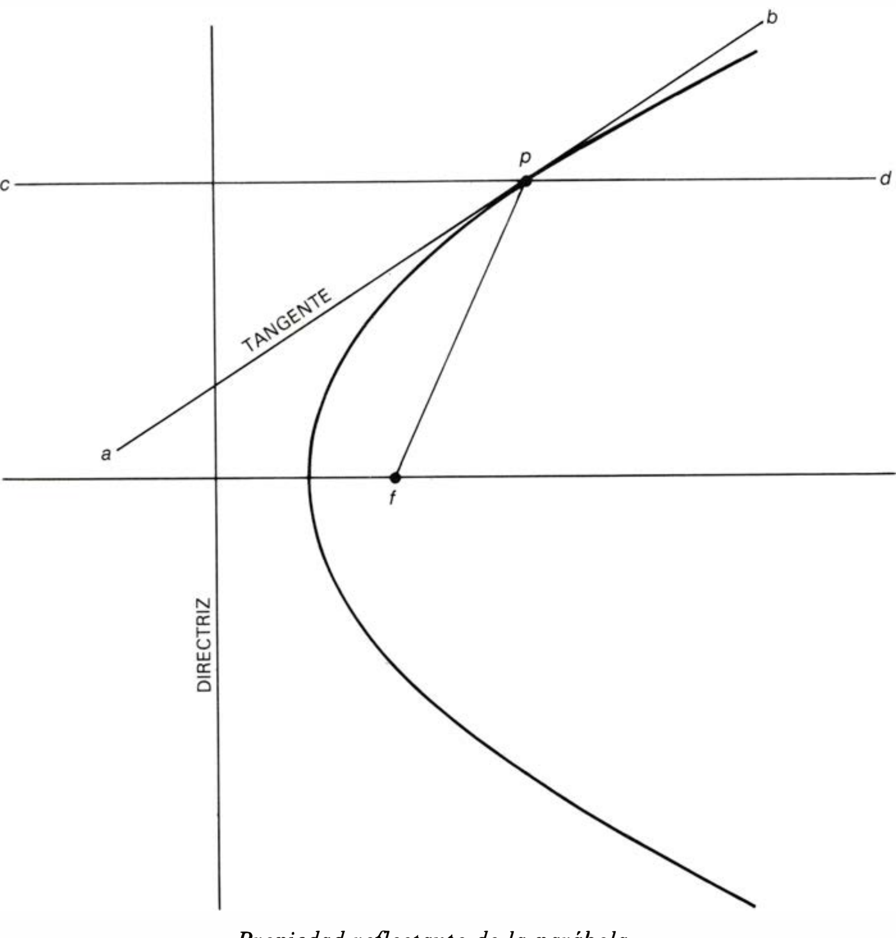
Al lanzar una piedra horizontalmente, su trayectoria se acerca mucho a una parábola, como ya fue demostrado por Galileo en 1609, pese a no publicar la demostración hasta casi 30 años más tarde. Podemos repetir uno de los experimentos de Galileo haciendo rodar sobre un plano inclinado una bola mojada en tinta, lanzada al bies. Recubriendo el plano con papel cuadriculado, la línea descrita por la bola al rodar sobre él nos permitirá ajustar una ecuación cuadrática que la describa. En realidad, la trayectoria de los proyectiles no es parabólica, debido, por una parte, a la redondez de la Tierra y, en proporción muchísimo mayor, a la resistencia del aire. En sus *Discorsi e dimostrazione matematiche intorno a due nuove scienze* (1638), Galileo examina ampliamente la influencia de la viscosidad del aire y de la redondez del globo terráqueo; hoy nos haría sonreír cuando dice no tomar en cuenta los efectos de la redondez de la Tierra, pues el alcance de los proyectiles militares "nunca excederá de cuatro millas".

La resistencia del aire al vuelo de la bala hace que la trayectoria de ésta se parezca tanto al perfil de un pecho que, hacia el final de su novela *The Naked and the Dead,* Norman Mailer hace trazar a uno de sus personajes, oficial del ejército, varias gráficas de la curva, mientras musita:

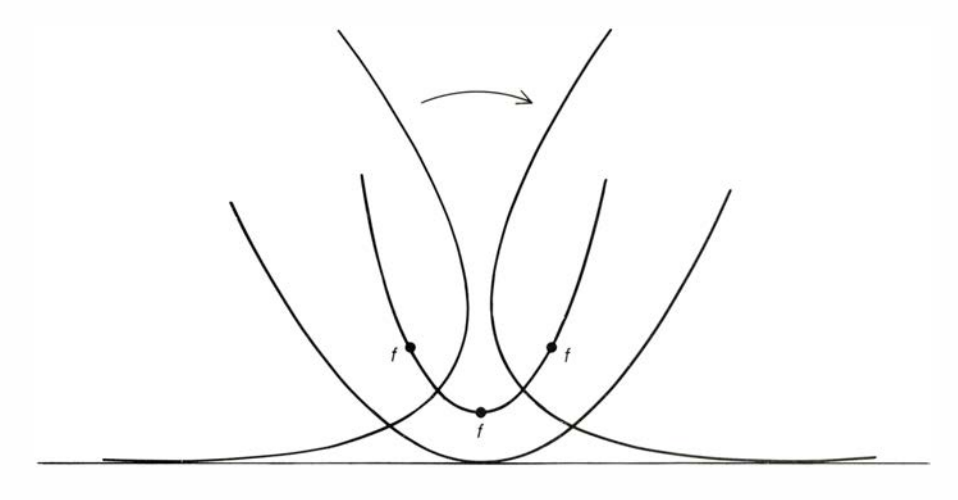
"Esta forma... es la curva fundamental del amor, según creo. Es la curva que siguen todas las potencias del hombre (si prescindimos de las fases de "meseta" debidas a la educación, a los correctivos contra la decadencia). Parece ser también la curva de la excitación y descarga sexual, que después de todo es el núcleo donde se concreta la vida. ¿Qué representa esta curva? *Es la trayectoria fundamental de cualquier pro-*

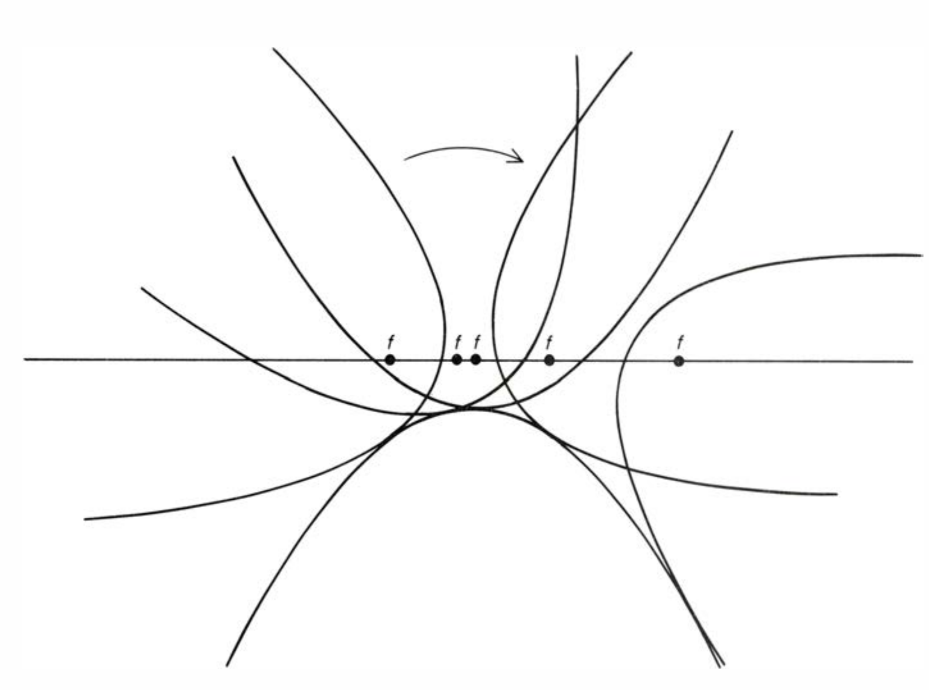


Las parábolas $y=5(x-1)^2+3$ **(arriba)** *e* $y=-1/8\,(x+1)^2+2$ *tienen forma idéntica*

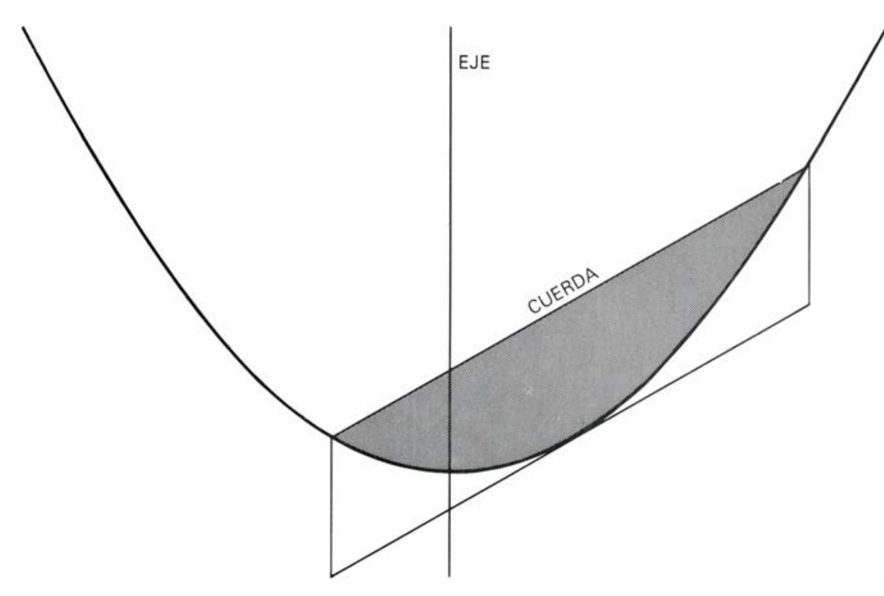


Propiedad reflectante de la parábola





Al hacer rodar parábolas se generan catenarias **(arriba)** *y líneas rectas* **(abajo)**



El segmento sombreado tiene 2/3 del área del paralelogramo

yectil, de una piedra, de una bala, de una flecha (y también de la flecha del anhelo de que hablaba Nietzsche), o de una granada de artillería. Tanto es la curva del proyectil mortal como del impulso vida-amor; pone de relieve la forma en que se manifiesta la existencia: la vida y la muerte son tan sólo distintos puntos de observación de una misma trayectoria. Corresponde a la vida lo que vemos y sentimos cuando cabalgamos sobre el proyectil; es el presente, el ver, el sentir, el percibir. Desde el observatorio de la muerte se ve la granada como un todo, se comprende su fin inexorable, el punto hacia el que ha sido destinado por leyes físicas inevitables desde el instante en que recibió su impulso primario y fue catapultada en el aire.

"Llevemos el razonamiento un paso más allá. Hay dos fuerzas que constriñen al proyectil a seguir su trayectoria. De no ser por ellas, su camino ascendería indefinidamente en línea recta, ↗. Dichas fuerzas son la gravedad y el rozamiento del aire, y sus efectos son proporcionales a los cuadrados de los tiempos; van haciéndose cada vez mayores, y en cierto sentido se alimentan de sí mismos. El proyectil aspira a avanzar por el camino ↗, la gravedad, a llevarlo hacia abajo, ↓, y la resistencia del aire actúa ←. Estas fuerzas parásitas crecen y crecen al transcurrir el tiempo, acentuando el declive de la trayectoria, reduciendo su alcance. Si sólo actuase la gravedad, el arco descrito sería simétrico



y es la resistencia del aire la que provoca la trágica curva.



"En su acepción más amplia, la gravedad ocuparía el sitio de la mortalidad (lo que asciende tiene que descender) y la resistencia del aire sería la oposición del ambiente... la masa inercial, o si se quiere, la inercia de las masas, a través de la cual la visión, el impulso de la cultura hacia lo alto, va perdiendo su filo, desgastando su agudeza, y es frenado y llevado finalmente a sacrificio prematuro."

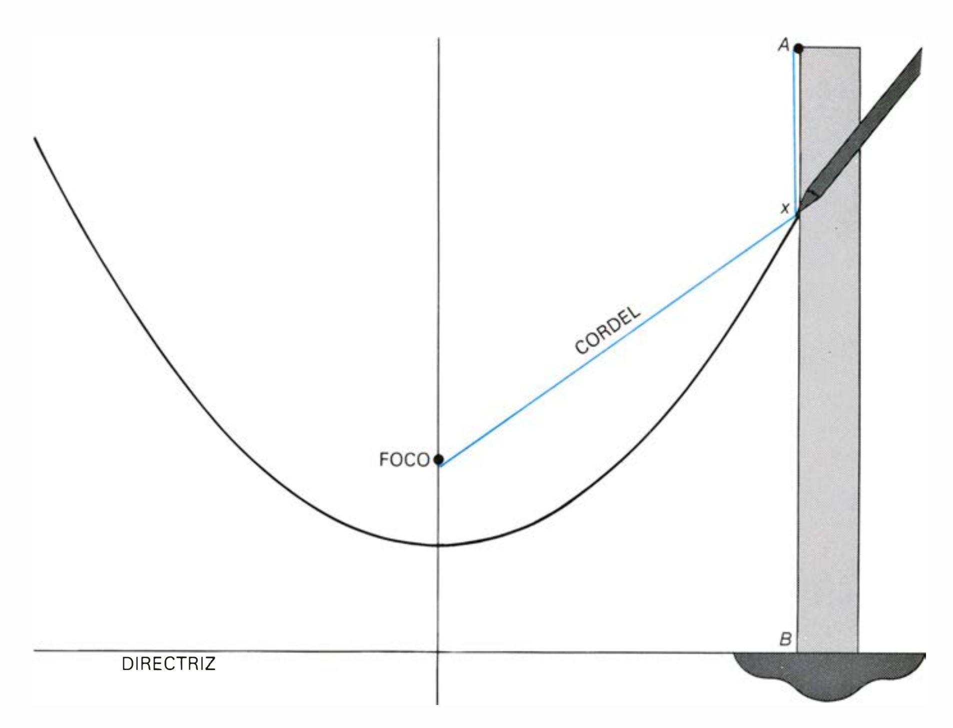
También un chorro de agua que mana de una manguera sigue en su vuelo una parábola casi perfecta. Si al regar el césped vamos lentamente reduciendo el ángulo del chorro, desde casi la vertical hasta la horizontal, el punto de altura máxima de cada una de las curvas parabólicas, de la serie que va describiéndose conforme decrecen los ángulos, va trazando una parábola en el

aire, en el mismo plano barrido por los chorros.

Pudiera ser que algunos cometas siguieran trayectorias parabólicas. Los cometas que periódicamente regresan al sistema solar recorren órbitas elípticas extraordinariamente excéntricas, pero (como ya hemos visto) cuanto más excéntrica es una elipse tanto más se asemeja a una parábola. Siendo la parábola una curva límite entre las familias de las elipses y de las hipérbolas, es casi imposible decir, basándonos en observaciones del cometa en las vecindades del Sol, si aquél está describiendo una elipse muy excéntrica (en cuyo caso el cometa retornará) o si su trayectoria es una parábola o una hipérbola (no retornando jamás). La trayectoria será parabólica cuando la velocidad del cometa sea igual a la velocidad de escape del sistema solar. Cuando su velocidad sea mayor, la trayectoria será hiperbólica, y si es menor, describirá una órbita elíptica.

Las notables aplicaciones de la parábola en problemas técnicos se inspiran, casi todas, en la propiedad de reflexión que vemos ilustrada en la figura inferior de la página 119. Tracemos una recta que salga del foco f e intercepte a la curva en un punto cualquiera p. Sea ab la tangente a la parábola por el punto p. Una recta cd, que pasando por p defina ángulos apf y bpd de igual amplitud, tendrá que ser ortogonal a la directriz. Se deduce entonces que, al imaginar la parábola como curva reflectante, todo rayo de luz que emane del foco se reflejará en la curva siguiendo una recta paralela al eje de aquélla.

Supongamos ahora que hacemos girar la parábola en torno a su eje, engendrando así la superficie denominada paraboloide de revolución. Los rayos de luz que partan del foco, tras chocar con la superficie del paraboloide, se reflejarán formando un haz paralelo al eje del paraboloide. Todos los faros y proyectores de luz se inspiran en este principio. Evidentemente, la propiedad puede aplicarse en sentido inverso: rayos luminosos paralelos, que incidan según la dirección del eje de un espejo parabólico cóncavo, serán concentrados en el foco. Este es el secreto de los telescopios reflectores, de los concentradores de energía solar y de las antenas de microondas. Siendo mucho más sencillo construir espejos parabólicos grandes que lentes transparentes de tamaño equiparable, todos los telescopios gigantes son hoy de tipo reflector. Otros dispositivos ópticos sirven para llevar la imagen hasta un ocular o placa fotográfica. Quizá de niño haya jugado usted a incendiar una hoja de papel concentrando en ella los rayos solares con una lente de cristal. Es igualmente fácil conseguirlo situando el papel en el foco de un espejo parabólico.



Construcción de Johannes Kepler con regla y cordel

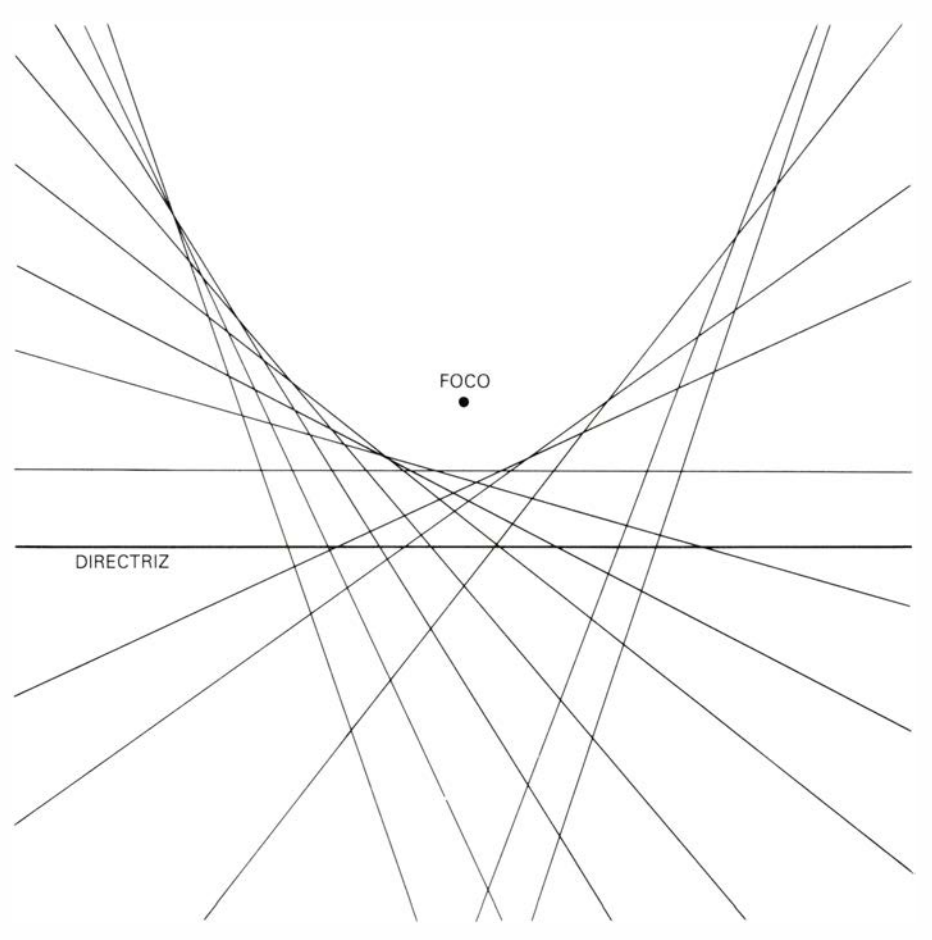
Al hacer girar un cubo que contiene agua, la superficie del líquido acaba formando un paraboloide; esta propiedad dio al físico R. W. Wood la idea de que quizá fuese posible construir un telescopio reflector haciendo girar un gran plato de mercurio, utilizando como espejo la superficie parabólica así creada. Wood llegó a construir un prototipo, pero tropezó con tantas dificultades para lograr una superficie lisa que abandonó la idea.

Imaginemos un paraboloide de base plana y ortogonal al eje, semejante a una colina bien redondeada. ¿Cómo calcular su volumen? Arquímedes encontró para este problema una solución de sencillez sorprendente: el volumen es, exactamente, 3/2 del volumen de un cono recto de igual base circular e idéntica altura.

Los cables de suspensión de los puentes colgantes trazan curvas muy aproximadamente parabólicas. La curva se deforma cuando el peso del puente no está uniformemente repartido o cuando el peso de los cables es grande con relación al peso del puente. En este último caso, la curva apenas se distingue de la llamada *catenaria* (del latín "catena", cadena). Galileo supuso erróneamente que al suspender una cadena por sus extremos resulta una parábola. Tuvieron que pasar decenios hasta descubrirse que no es así, y que la catenaria es una curva no algebraica, pues su ecuación se basa en el número e, que es trascendente.

Entre parábolas y catenarias existe una relación muy curiosa y poco conocida. Al hacer rodar una parábola sobre una recta, como vemos en la ilustración superior de la página adyacente, el "lugar geométrico de los focos" es una catenaria perfecta. Quizá más sorprendente (y desde luego, más fácil de probar) sea lo que sucede cuando dos parábolas de igual tamaño se sitúan vértice contra vértice y se hace rodar una sobre la otra, como se muestra en la parte baja de la ilustración superior de la página contigua. ¡El foco de la "ruleta" describe una línea recta, y su vértice, una cisoide!

Uno de los primeros problemas relativos a parábolas fue la "cuadratura" del área limitada por un arco de curva y el segmento rectilíneo que une sus extremos, es decir, de la región sombreada en la ilustración inferior de la página adyacente. El primero en resolver este problema fue Arquímedes, en su famoso tratado *Cuadratura de la parábola*. Valiéndose de un ingenioso paso al límite, que muchos siglos después sería inspiración del cálculo integral, pudo demostrar que, circunscribiendo a la región el paralelogramo que vemos en la figura, el área del segmento parabólico es 2/3 del área del paralelogramo. (Inicialmente, Arquímedes determinó empíricamente esta relación, comparando el peso del palalelogramo con el peso del segmento parabólico.) Arquímedes se sirvió también de la parábola para dar una elegante construcción del heptágono regular. Geómetras anteriores se habían valido ya de parábolas para el clásico problema



Parábola construida por dobleces

de duplicar el cubo: construir un cubo de volumen doble al de otro cubo dado.

Hay muchas técnicas para trazar parábolas sin tener que representar miríadas de puntos en la hoja de papel. Quizá la más sencilla sea la que se vale de una escuadra en T y un cordel. Uno de los extremos del cordel se sujeta al extremo del brazo de la T, como muestra la ilustración de la página 121; el otro extremo se fija al foco de la parábola. El hilo debe tener precisamente la longitud AB. La punta del lápiz, situada en x, se apoya contra el brazo de la escuadra y mantiene el hilo tirante. Conforme la T se desliza hacia la derecha a lo largo de la directriz, el lápiz va ascendiendo por el brazo de la T y traza la rama derecha de la parábola. Copiando la figura por simetría, o repitiendo el trazado hacia la izquierda, se obtiene la otra rama. El método fue inventado, o posiblemente, reinventado, por Kepler. Es evidente que en lugar de la escuadra en T puede usarse una escuadra o cartabón triangular ordinarios, haciéndola deslizar por el borde de una regla. Cuesta poco verificar que, al ser el hilo de longitud constante, queda asegurado que cualquier punto de la curva se encuentra a igual distancia del foco y de la directriz.

Plegando una hoja de papel es aún más fácil lograr preciosas parábolas. Basta situar el foco en un punto cualquiera de una hoja de papel vegetal, dibujar una directriz con la regla, e ir doblando la hoja muchas veces, haciendo siempre que, al doblar, la directriz pase por el foco. La arista creada por cada doblez será tangente a una misma parábola, que aparece ahora como envolvente de esta familia de rectas, según sugiere la ilustración de esta misma página. Si el papel es opaco, puede usarse uno de sus lados como directriz; en cada pliegue este lado debe pasar por el foco.

La familiaridad con las propiedades de la parábola puede ayudar a resolver rápidamente problemas algebraicos. Tomemos, por ejemplo, el sistema de ecuaciones siguiente, cuyos términos independientes son las puntuaciones de la suerte al jugar a los dados:

$$x^2 + y = 7$$
$$x + y^2 = 11$$

No cuesta mucho descubrir que si x vale 2 e y vale 3 se verifican ambas ecuaciones. Plantearemos dos cuestiones:

1. ¿Existirán otras soluciones enteras (donde x e y sean ambos números enteros, positivos o negativos)?

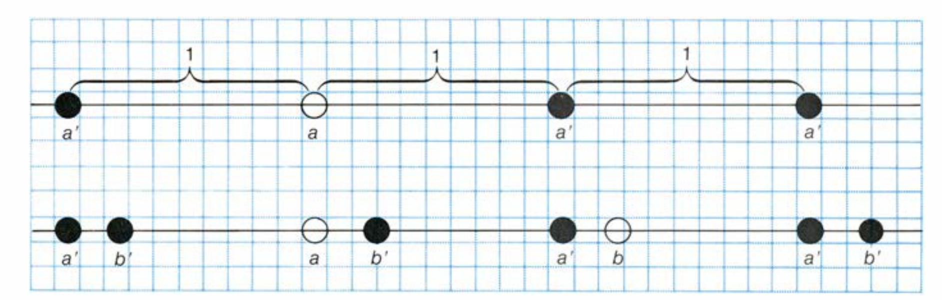
2. ¿Cuántas soluciones habrá en total?

Un problema más difícil, que se resuelve gracias a una parábola, es debido a Ronald L. Graham, de Bell Laboratories, quien lo ha propuesto y también resuelto. Inédito hasta ahora, se publica aquí por primera vez. Supongamos que se disponga de una colección infinita de discos idénticos cuyo diámetro sea menor que 1/2 de cierta unidad, por ejemplo, diámetro 1/10. ¿Será posible colocarlos todos en el plano sin que se solapen unos con otros, y sin que nunca la distancia entre pares de puntos tomados de los discos venga dada por un número entero?

Como el diámetro de cada disco es solamente 1/10, ningún par de puntos distintos de un mismo disco puede distar un número entero de unidades, pero sí es concebible que, disponiendo sagazmente los discos con sus centros en una recta, ningún punto de ninguno de los discos se encuentre a distancia entera de ningún otro punto del mismo tipo. No es difícil demostrar que esto es imposible; aún más, toda disposición de la infinidad de discos en una recta dará lugar a una infinidad de pares de discos en los cuales una infinidad de pares de puntos (exactamente todos los de la línea) se encontrarán separados por distancias enteras.

Para comprenderlo, imaginemos situado un disco de diámetro 1/10 con su centro sobre la recta, representado por un círculo en blanco a en lo alto de la ilustración superior de la página 124. Los círculos oscuros (marcados a') están espaciados por igual, con distancia centro a centro unitaria, y se extienden, indefinidamente hasta el infinito, en ambas direcciones. Es evidente que no podrá colocarse un segundo disco en ningún lugar donde toque o cabalgue sobre alguno de los círculos oscuros, pues tal disco contendría entonces puntos (y más aún, puntos de la recta) situados a distancias enteras de puntos del disco a. (Estamos suponiendo que los puntos de la circunferencia del disco pertenecen también a él.)

Evidentemente, es posible situar un segundo disco entre cualquier par de círculos adyacentes, con tal de que no toque ni recubra ninguno de ellos. Por ejemplo, podríamos situar un segundo disco b como se muestra en la parte baja de la ilustración superior de la página 124. La recta pasa inmediatamente a disponer de otra familia infinita de círculos oscuros (marcados b') separados una unidad de centro a centro, indicando así que en tales lugares no podrá colocarse un tercer disco que los toque o recubra; lo mismo puede decirse para discos sucesivos. Dado que será



Demostración de la imposibilidad de la conjetura con líneas rectas

imposible situar más de ocho discos disjuntos en ninguno de los intervalos definidos por dos discos consecutivos de la primera familia, vemos que será imposible situar sobre la recta más de nueve discos. Un décimo, no importa dónde sea situado, contendrá una infinidad de puntos separados por número entero de unidades de puntos de alguno de los nueve discos ya colocados.

La demostración se generaliza sin dificultad a todos los discos de diámetro menor que 1. Si el denominador del diámetro es número entero, basta restar 1 para obtener el máximo de discos admisibles. Si el denominador no fuese entero, se redondearía al entero más cercano. Por supuesto, no pueden usarse discos de diámetro unidad o mayor. Solamente podrá situarse en la recta un disco de diámetro $1/2$ o $1/\sqrt{2}$; dos, de diámetro $1/3$; tan sólo tres de diámetro $1/4$ o de diámetro $1/\pi$; cuatro, de diámetro 1/4'5, y así sucesivamente.

Aunque el problema es insoluble con líneas rectas, sí puede resolverse con parábolas. En mi próxima sección presentaré la solución de Graham, y resolveré también las cuestiones precedentes.

El mes de agosto se les pedía a los lectores que situasen 20 damas sobre un tablero de ajedrez, de forma que cada una amenace a otras cuatro. Damos una solución en la ilustración inferior de esta misma página. Se desconoce si podrán situarse en el tablero más de 20 damas que cumplan esta misma condición.

La solución más conocida del problema de las damas de ajedrez

Una carta de D. Weeks me señala un error en mi sección sobre congruencias, del mes de abril. Comentando allí el teorema de Wilson dije que uno de sus contemporáneos (Edward Waring) afirmó que el teorema nunca sería demostrado. Tal apostilla, originariamente en latín, en su traducción correcta dice que el teorema será muy difícil de demostrar. Por otra parte, al decir que en 1582 el equinoccio vernal se iba aproximando al invierno quise decir que la fecha en que tradicionalmente se suponía ocurría el equinoccio iba deslizándose hacia el invierno, y no, claro está, que así le sucediese al equinoccio astronómico.

Y cuando dije que la fecha "más frecuente" de la Pascua era el 19 de abril, estaba basándome en datos relativos a un margen temporal relativamente corto. La cosa no es tan sencilla. Varios lectores se valieron de ordenadores para revisar otros intervalos temporales, y obtuvieron resultados variables. James LaBossiere, Richard Morrish y otros, examinando el período cubierto por el algoritmo de Thomas H. O'Beirne (desde 1900 hasta 2099), descubrieron que el 31 de marzo, el 12 de abril y el 15 de abril quedaban empatados como fechas más probables, sin que el 22 de marzo apareciera para nada en todo este período.

Thomas L. Lincoln, L. V. Larsen y otros varios investigaron distintos intervalos, con resultados mutuamente contradictorios. Lincoln opina que, a largo plazo, de seguir en vigor las reglas que actualmente determinan la Pascua, el 19 de abril llegará a ser la fecha más corriente, pero tal afirmación no ha sido confirmada. Trataré de dar cuenta de nuevos progresos sobre este problema al parecer no resuelto todavía de una manera concluyente.

Taller y laboratorio

Imágenes de interferencia en un espejo polvoriento

Jearl Walker

Contemplando un espejo empañado o cubierto de polvo, mientras se sostiene una pequeña luz delante de la cara, pueden observarse dos bellos casos de interferencia óptica; para ello puede servir un fósforo encendido. Lo que se ve entonces son círculos de colores centrados en la imagen de la luz. Solapado a éstos existe otro conjunto de círculos y colores de interferencia, pero puede que no esté centrado alrededor del mismo punto. Esta otra figura no se ve tan fácilmente, porque, para ello, el ojo y el manantial luminoso deben encontrarse los dos en la misma perpendicular al espejo. No obstante, corrigiendo cuidadosamente las posiciones del manantial y del ojo propio, se consigue que el segundo conjunto de círculos coincida con el primero. En ciertas circunstancias, los colores que exhiben estas figuras son asombrosos.

Ambas figuras de interferencia requieren un espejo con la superficie reflectora en el reverso, caso en el que se encuentran los espejos ordinarios. La figura que aparece siempre centrada en la imagen reflejada del manantial luminoso suele conocerse como figura de Fraunhofer; ha sido ya objeto de estudio exhaustivo, no sólo porque depara una explicación excelente de la difracción óptica, sino, además, porque esa interferencia es responsable del anillo coloreado que a veces se observa rodeando la luna o el sol.

La otra imagen interferencial que puede observarse en los espejos empañados o polvorientos se investigó intensamente a lo largo del siglo XIX, pero desde entonces ha sido prácticamente olvidada. Este mes voy a comenzar hablándoles de esa interferencia menos conocida; luego regresaremos a la figura de Fraunhofer y a los anillos lunares y solares.

Parece cierto que fue Isaac Newton el primero en abordar las figuras de interferencia coloreadas que se forman en los espejos cubiertos de polvo. Para ello dirigía un rayo de luz solar sobre un orificio angosto practicado en una pantalla colocada frente a un espejo cóncavo semiesférico. La luz atravesaba el orificio, se reflejaba en el espejo y regresaba al orificio que estaba situado en el centro de curvatura del espejo. Entonces, la figura de interferencia aparecía en la pantalla rodeando el orificio.

No deja de ser irónico que, en la totalidad de sus investigaciones en el terreno de la óptica, Newton desestimara la posibilidad de que la luz estuviera constituida por ondas. Así, aferrado a la idea de que la luz se componía de partículas, no consiguió explicar ninguna de las manifestaciones de interferencia luminosa que él mismo estudió, incluida la que observaba en los espejos polvorientos. El conocimiento correcto de esta cuestión no surgió hasta que Thomas Young probó de manera convincente la naturaleza ondulatoria de la luz y pudo explicar, rápidamente, los colores que se veían en los espejos polvorientos de Newton. William Whewell, en Inglaterra, y Lambert Adolphe Jacques Quételet, en Bélgica, siguieron investigando en torno a esas figuras; y emplearon espejos planos, razón por la cual reciben hoy el nombre de figuras de interferencia de Whewell y Quételet.

Casi todas las partículas (polvo ambiental o de yeso, polvos de licopodio) que se depositen sobre un espejo pueden crear la figura de Whewell-Quételet. También sirve la humedad que se condensa cuando empañamos, con nuestro aliento, el espejo. Otra posibilidad es crear una capa delgada de glóbulos de leche seca. Para ello, se mezcla una parte de leche con tres de agua y se recubre la superficie anterior del espejo. Luego, se inclina éste para que el líquido en exceso se escurra al exterior y se deja secar el resto. Los glóbulos de leche que hayan quedado harán el papel de partículas de polvo. Por otro lado, no es necesario que la sustancia pulverulenta se halle en contacto directo con el espejo; Sir John Herschel ponía de manifiesto la figura de interferencia lanzando al aire polvos de peluca delante de un espejo.

La figura de interferencia se debe a la dispersión de la luz por las partículas en contacto con el espejo, o que se encuentran en su proximidad. Para simplificar, voy a suponer que el observador y el manantial luminoso se hallan en la misma perpendicular al espejo y que la luz procedente del manantial se compone de una sola longitud de onda. De la interferencia que se produce con luz blanca me ocuparé después y les explicaré cómo se descomponen los colores.

Consideremos dos rayos que llegan al espejo procedentes del manantial, formando parte de un frente de onda. Estos rayos son coherentes (o sea, las ondas luminosas de ambos están en fase) y casi paralelos, pues, aunque el manantial luminoso sea incoherente (salvo que sea un láser), éste sí emite ondas coherentes a pequeñas ráfagas de corta longitud.

Imaginemos que una partícula dispersa uno de los rayos, lo envía a la superficie posterior del espejo y allí el rayo se refleja y retrocede hacia el observador. Supongamos que el otro rayo penetra en el espejo sin haber sido dispersado por la partícula, se refleja y lo dispersa entonces la partícula en cuestión dirigiéndolo hacia el observador. Inicialmente, los rayos estaban en fase, pero ahora tendrán una diferencia de fase que dependerá de los trayectos respectivamente recorridos. Si estos trayectos tienen la misma longitud, los rayos seguirán en fase, se interferirán constructivamente y el observador verá una luz bastante intensa procedente del punto del espejo ocupado por la partícula. Cuando el observador y el manantial luminoso estén en la misma perpendicular al espejo, tal situación corresponderá a una partícula situada directamente frente al observador, o sea, situada al pie de la perpendicular.

Una partícula ligeramente separada del pie de la perpendicular enviará al observador rayos con una diferencia de fase distinta y, si la diferencia entre los trayectos produce una diferencia de fase de media longitud de onda, los rayos se interferirán destructivamente y el observador verá un punto oscuro. Más aún, originado por las partículas de polvo situadas en un círculo centrado en el pie de la perpendicular, cuya posición sea la precisa para causar interferencia destructiva, el observador verá, además, un anillo oscuro rodeando el pie de la perpendicular; se trata del anillo oscuro más cercano al centro de la figura de interferencia.

Las partículas un poco más alejadas del centro dispersan las parejas de rayos que llegan al observador con una diferencia de fase de una longitud de onda completa. Aquí el resultado será interferencia constructiva y un anillo luminoso. Algo más separados del pie de la perpendicular se verán más anillos claros y oscuros, causados por diferen-

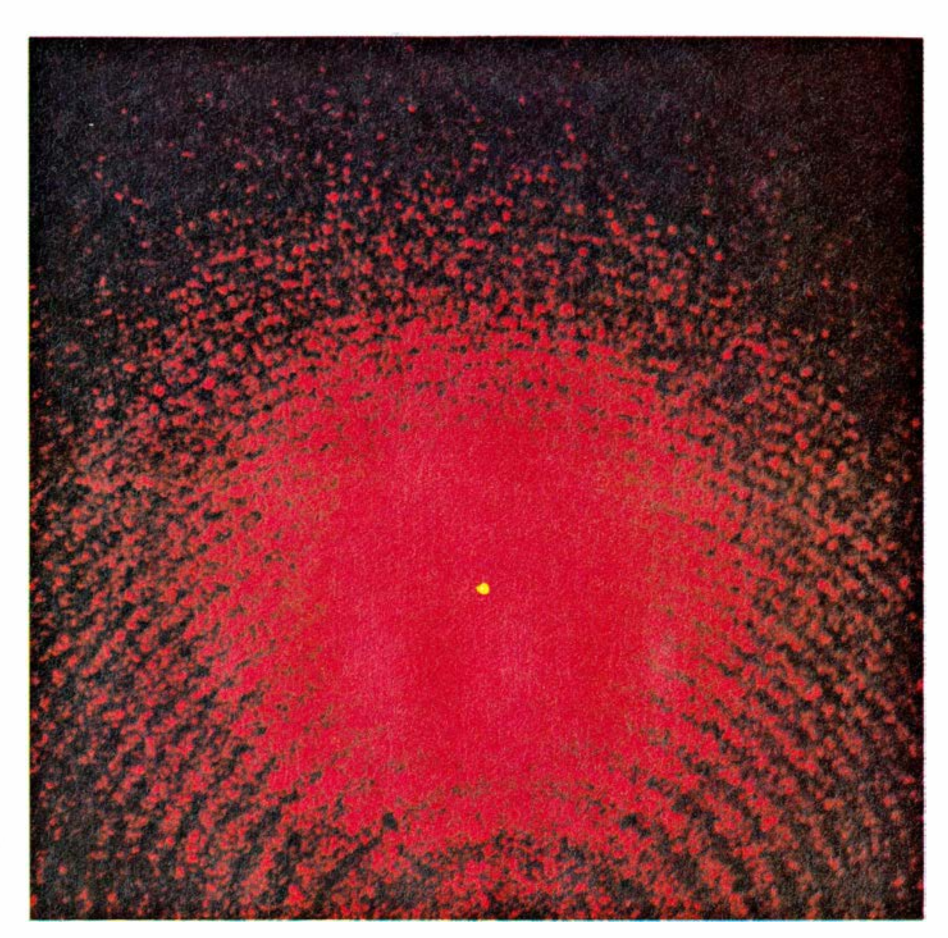
cias progresivamente mayores entre los trayectos de las parejas de rayos luminosos.

Las figuras de interferencia suelen designarse en concordancia con la diferencia de fase entre los rayos que participan en ellas. Así, la primera posición de luminosidad, para la cual la diferencia de fase es cero, se designa $n = 0$ y recibe el nombre de franja de interferencia constructiva de orden cero. Tal franja, en la disposición geométrica de observador y manantial luminoso que les he descrito, se encuentra al pie de la perpendicular. La franja luminosa siguiente, correspondiente a la diferencia de fase de una longitud de onda completa, se designa $n = 1$ y se llama franja de interferencia constructiva de orden uno; esta franja es el anillo luminoso situado en el centro de la figura de interferencia.

En los ejemplos clásicos de interferencia, los índices numéricos de las franjas luminosas aumentan con la distancia al centro de la figura. El caso de la imagen interferencial creada en un espejo cubierto de polvo constituye una excepción, al menos en principio. Aquí, si pudieran observarse los anillos luminosos aún más separados del centro, el orden de éstos comenzaría a decrecer. El motivo de ese comportamiento estriba en que, cuando dos rayos regresan al observador desde un punto relativamente alejado de la perpendicular, habrán viajado distancias casi iguales y, por consiguiente, su diferencia de fase será menor que la de otra pareja de rayos que se dispersen en una partícula algo más próxima a la perpendicular. Es una pena que los anillos en los que se invierte el orden numérico no puedan observarse en los espejos ordinarios.

Además, la anchura de los anillos varía con su distancia al centro de la figura, haciéndose más estrechos al principio y ensanchándose después (demasiado lejos del centro para que normalmente se vea). La anchura del anillo perteneciente a un orden de interferencia determinado depende de varios factores. Así, un anillo es tanto mayor cuanto mayores son la longitud de onda y el índice de refracción del vidrio del espejo y menor el espesor de éste. La anchura depende asimismo de la relación entre la distancia de la fuente luminosa al espejo y la distancia del observador al espejo.

¿Por qué han de dispersarse en la misma partícula una pareja de rayos para que participen en la creación de una figura de interferencia? La respuesta es que el rayo procedente de una partícula puede interferir con el procedente de otra, pero esta interferencia no contribuye a formar una figura estruc-



*Figuras de Fraunhofer (**arriba**) y de Whewell y Quételet (**abajo**) con luz láser*



Figura de Fraunhofer con polvos de licopodio y luz blanca

turada. Efectivamente, la dispersión de la luz causada por una partícula produce un cambio en la fase de la luz que depende del tamaño y forma de la partícula; entonces, para dos rayos dispersados por la misma partícula, el cambio de fase es el mismo. Así pues, la diferencia de fase depende únicamente de las relaciones geométricas entre el manantial luminoso, el observador y la partícula.

El tipo de dispersión originada por una partícula en contacto con el espejo depende de su tamaño. Las mayores que 100 micrometros dispersan la luz por reflexión y por difracción a la vez. Las más pequeñas, como los polvos de licopodio, dispersan la luz mediante un proceso más complicado, en el que intervienen ondas superficiales. En la teoría de la dispersión de Mie se encuentran los fundamentos de la explicación pormenorizada de este tipo de dispersión. Ocurre en este caso que, si las partículas están muy apiñadas en el espejo, adquiere importancia la interferencia entre los rayos que se difractan en los espacios de separación entre partículas; no vamos a entrar aquí en tales complejidades.

Con luz blanca, los anillos aparecen coloreados; y así, cada orden de interferencia presenta una secuencia de colores que van del azul al rojo, ordenados hacia fuera desde el centro de la figura. Esta separación de los colores se debe a que la diferencia de fase entre cada par de rayos participantes en la interferencia depende de la longitud de onda.

Supongamos que una partícula ligeramente separada del centro de la figura produzca interferencia constructiva de luz azul en el anillo claro más interno ($n = 1$). Dos rayos que contribuyan a este anillo recorrerán sendos trayectos cuya diferencia equivaldrá a una longitud de onda completa de luz azul. Si pretendemos que, con luz roja (cuya longitud de onda es mayor), se tenga la misma diferencia de fase, los dos rayos deberán dispersarse en una partícula un poquito más alejada del centro, para que la diferencia entre los trayectos equivalga a una longitud de onda completa de luz roja. De aquí que el interior del anillo sea azul y su exterior rojo; entre ambos se ven colores intermedios.

La figura descrita es la interferencial del conjunto visible cuando observador y manantial luminoso se hallan en la misma perpendicular. Si el observador se desplaza, el centro de la figura se separa de la imagen directamente reflejada del manantial luminoso y se presenta claro u oscuro, según cuánto se desplace el observador. A la vez que el observador se mueve, del centro parecen surgir nuevos anillos; y los más estrechos, correspondientes a órdenes de interferencia más elevados, pueden verse cercanos a la imagen reflejada del manantial. El centro de la figura termina por separarse tanto que desaparece y, si el observador se desplaza lateralmente lo suficiente, toda la figura se oscurece y se desvanece.

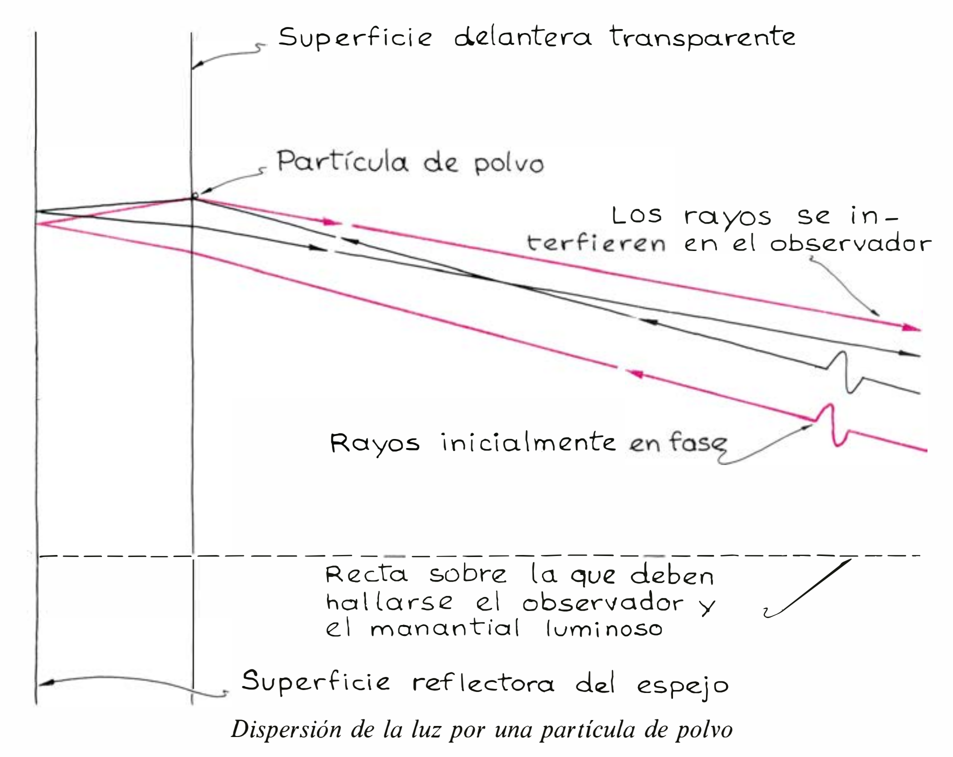
Para estudiar la estructura de las franjas de orden superior se empleó, de espejo, una lámina de mica delgada. Esta lámina, que se arranca de una placa más gruesa, se recubre por una de sus caras para hacerla reflectora. De ese modo puede verse una figura de interferencia, aun cuando el manantial luminoso y el observador estén muy desalineados.

La figura de interferencia también puede observarse en los espejos cuyo recubrimiento reflector lo tengan en la cara anterior. Para ello se empolva levemente una hoja de vidrio limpia y se coloca ante el espejo con su cara empolvada enfrentada a la superficie reflectora. Como separadores, entre el espejo y el vidrio se intercalan tiras de mica o papel. (Sin esta separación, toda pareja de rayos que producirían de suyo interferencia recorrerían la misma distancia, independientemente de la ubicación de la partícula de polvo en el vidrio, y no habría figura de interferencia.)

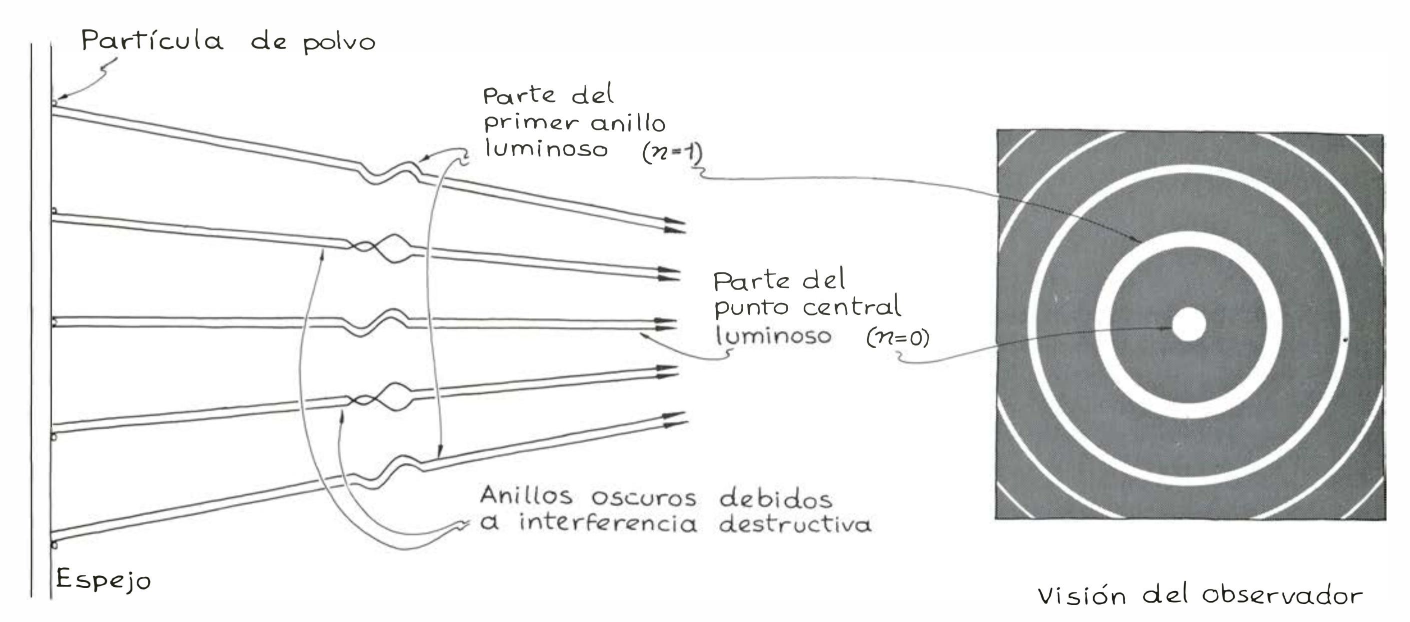
Por mi parte, para experimentar con estas interferencias, empleé una pequeña llama que sostuve entre un ojo y un espejo que mantenía empañado con mi aliento. El reverso del espejo me servía de superficie reflectora. Apagando las luces de la habitación pude ver tanto la figura de Fraunhofer como la de Whewell-Quételet, y observé que, a medida que me acercaba la llama al ojo, los anillos de la figura de Whewell-Quételet se hacían menos curvos y, con la llama junto al ojo, eran líneas casi rectas. Cuando un amigo sostuvo la llama detrás de mí (aunque situado de modo que no interceptara toda la luz con la cabeza), los anillos se curvaron.

Seguidamente, rocié el espejo con una leve capa de polvos de licopodio, agitándolo suavemente para conseguir una distribución lo más uniforme posible. (Si la capa es demasiado espesa, no se ve figura alguna.) Iluminé el espejo con una lámpara de incandescencia o una de sodio, colocadas a 10 metros de distancia para que se comportaran como manantiales puntuales. (También hubiera podido intercalar un orificio angosto entre un manantial más próximo al espejo y éste.) Los anillos de Fraunhofer se evidenciaron inmediatamente, pero me costó mucho más encontrar las franjas de Whewell-Quételet; tanto, que hube de poner la cabeza casi en la trayectoria del haz luminoso para lograr verlas. Un láser de helio-neon da los mismos resultados.

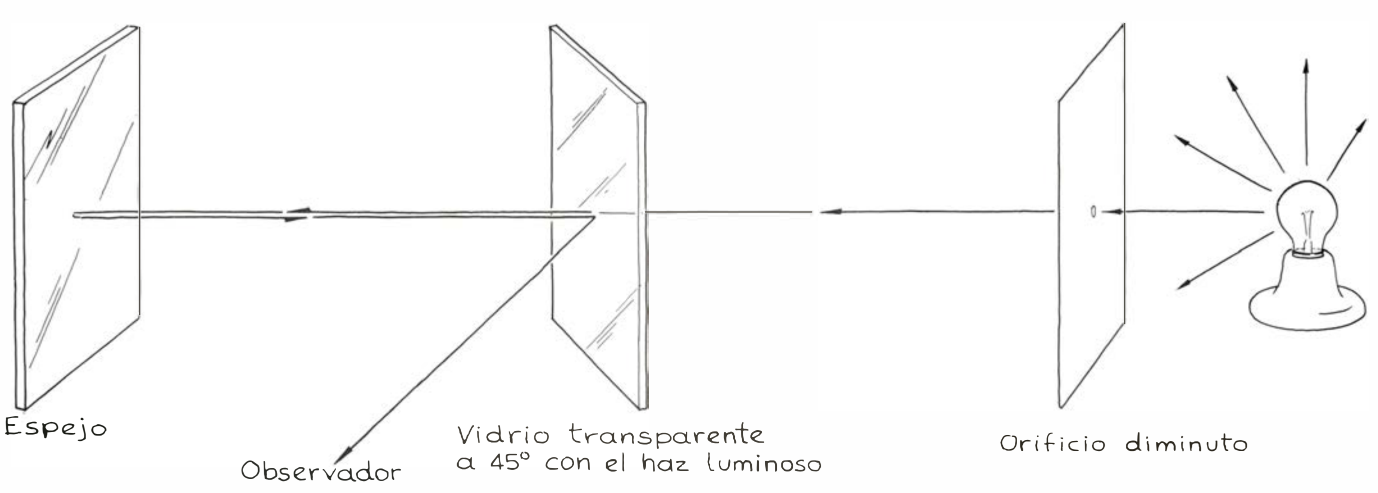
Para observar la figura de Whewell y Quételet más cómodamente coloqué una hoja de vidrio limpia en la trayectoria de la luz que se dirigía al espejo, dotándola del ángulo necesario para que proyectase un reflejo del espejo perpendicular al haz. Mirando al vidrio, me era posible ver en el espejo ambas figuras de interferencia. Cuando ajusté cuidadosamente los ángulos del espejo y del vidrio limpio, pude contemplar el



Dispersión de la luz por una partícula de polvo



Formación de una figura de Whewell y Quételet



Dispositivo para observar las figuras

espejo como si hubiera estado alineado con el manantial.

Hace 100 años, en Alemania, Eugen Lommel empleó por primera vez un dispositivo como el descrito para proyectar sobre una pantalla una imagen de la interferencia mediante una lente convexa. Este dispositivo sirve también para fotografiar la figura de interferencia. Basta con ajustar el objetivo de la cámara a una distancia focal grande, ya que, en principio, los rayos reflejados en el espejo convergen para dar una imagen real de la figura de Whewell-Quételet sólo a una distancia considerable del espejo. Con el objetivo en posición, la imagen se enfoca sobre la película.

Si usamos un láser como manantial luminoso, hay que andar con suma cautela con la luz reflejada en el espejo y en el espejo intercalado en el haz; ambos reflejos pueden ser suficientemente intensos para dañar la vista.

El otro fenómeno interferencial, que origina la figura de Fraunhofer, se debe a la difracción de la luz por las partículas de polvo depositadas sobre un espejo con el recubrimiento reflector en el reverso. Para explicarles lo que se observa les describiré primero una manifestación más sencilla de difracción por pequeñas partículas. Si, a través de una ventana empañada, miramos hacia un manantial luminoso pequeño, la imagen de éste aparece rodeada por anillos de interferencia concéntricos; se trata de anillos de Fraunhofer resultantes de la difracción de los rayos luminosos en las gotitas de agua de la ventana.

Cuando un haz de luz da con una partícula de pequeño tamaño, como una mota de polvo o una gotícula de agua, la luz se difracta alrededor del contorno de la partícula. Los rayos así difractados crean una imagen de interferencia al interceptarlos en una pantalla, tras separarse de la partícula. Si esta partícula presenta a la luz una sección aproximadamente circular, la figura de interferencia se hallará constituida de anillos luminosos concéntricos, que rodean a un punto central relativamente intenso, entremezclados con anillos relativamente oscuros.

Los anillos luminosos se deben a los rayos difractados en la partícula que se interfieren constructivamente; los oscuros se deben a interferencia destructiva. La clase de la interferencia depende de la geometría de los rayos, pues es la diferencia entre trayectos recorridos lo que produce la diferencia de fase entre aquéllos.

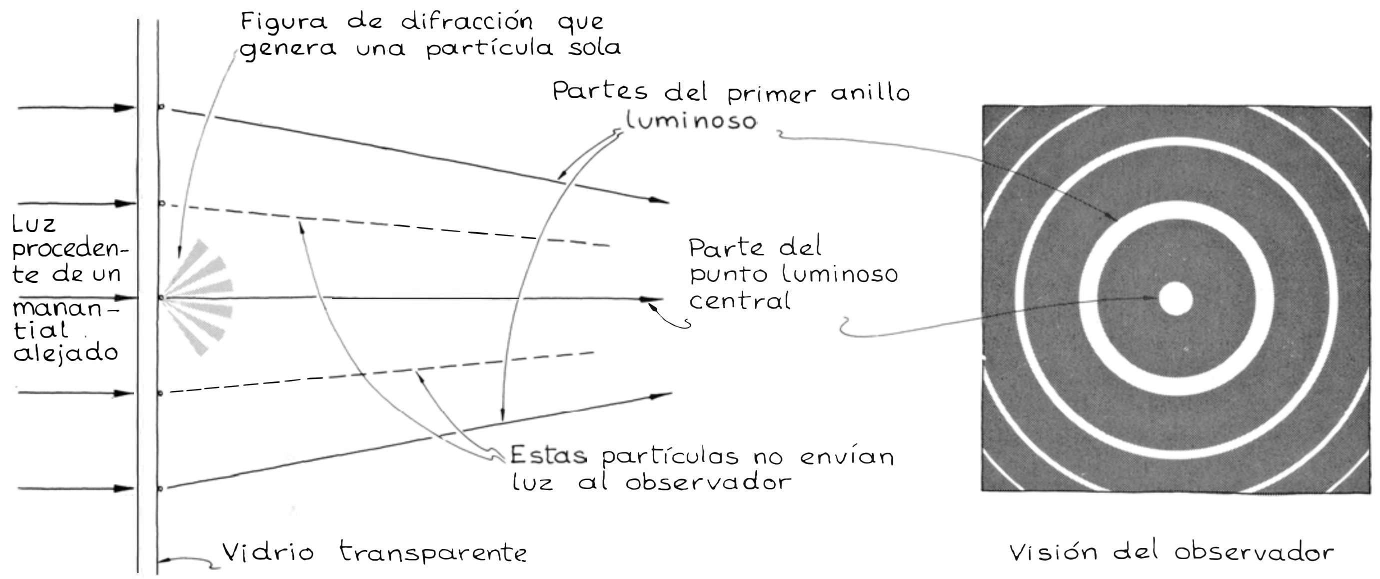
Al centro de la figura todos los rayos llegan en fase, pues todos han recorrido el mismo trayecto. Los rayos que llegan ligeramente separados del centro presentan unas diferencias en sus trayectos que dan por resultado una diferencia de fase de media longitud de onda; estos rayos se interfieren destructivamente.

Como dicha situación es la misma en todos los puntos de un círculo que rodee al punto central de la figura, en torno a éste aparecerá una banda oscura. El anillo siguiente es luminoso porque en él los rayos se interfieren constructivamente.

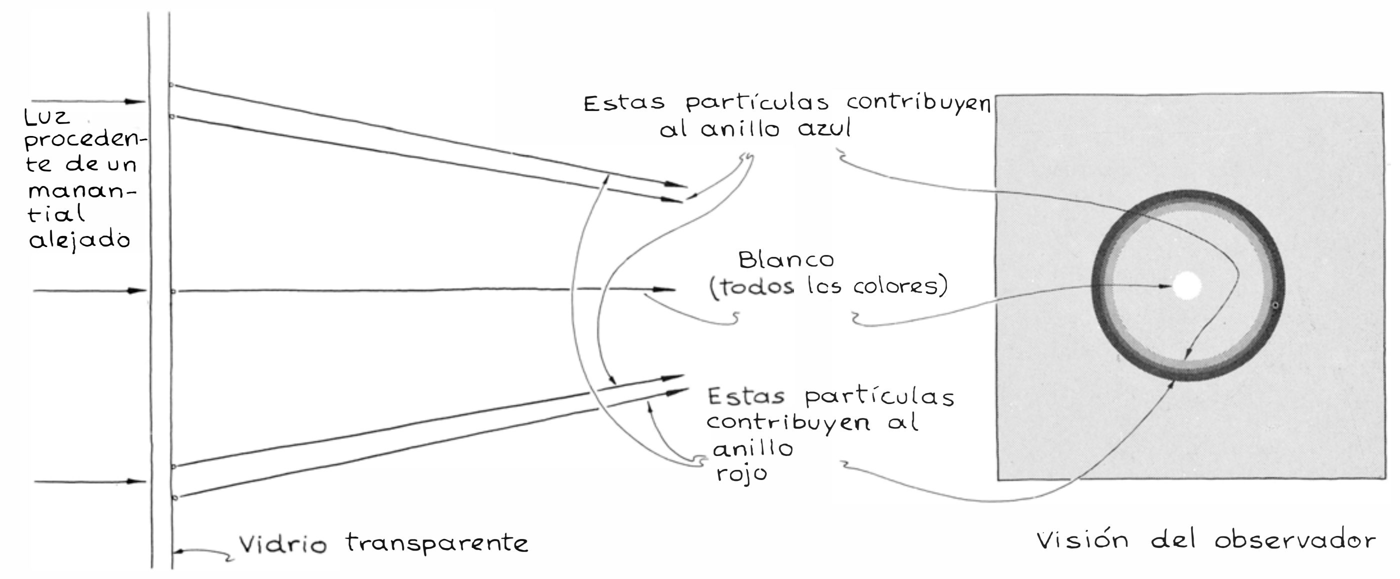
Descubriremos una figura similar si miramos a través de una hoja de vidrio empañada o cubierta de polvo, con tal que el manantial luminoso sea suficientemente pequeño, o esté alejado lo necesario para que ocupe un ángulo reducido en nuestro campo visual. No contemplaremos entonces la figura de difracción completa producida por cada partícula depositada en el vidrio, sino una figura compuesta por fragmentos de las figuras debidas a todas las partículas. Así, las partículas situadas en la recta que une el observador con el manantial luminoso envían los centros brillantes de sus figuras; las partículas ligeramente separadas de dicha recta envían una parte de su anillo oscuro más interno. Como resultado, vuelve a aparecer una figura de anillos alternativamente claros y oscuros. Además, las partículas no sólo difractan la luz para dar la figura de anillos de Fraunhofer, sino que también separan la luz blanca en sus colores componentes, por cuya razón los anillos se ven coloreados.

El anillo de Fraunhofer visible en torno de la luna o del sol se llama corona. (Término que no debe confundirse con la corona solar, que es la atmósfera luminosa que envuelve al Sol y que se hace visible durante los eclipses del planeta; tampoco debe confundirse con el halo, solar o lunar, que aparece como círculo en torno al astro, pero más separado de éste que la corona, y que se debe a la refracción, no a la difracción, de la luz en minúsculos cristales de hielo presentes en nubes de determinado tipo.) Las partículas responsables de este anillo son minúsculas gotas de agua o cristales de hielo que se encuentran en las nubes diáfanas situadas entre el observador y la luna o el sol.

Cuando la condensación me empaña las gafas o las ventanillas del automóvil se me ofrece la oportunidad de contemplar una corona similar. Entonces, las farolas lejanas y los faros de los automóviles aparecen rodeados de círculos de colores. A veces, también veo anillos de Fraunhofer cuando unos faros lejanos iluminan el espejo retrovisor de mi automóvil, que no es modelo de limpieza. (Hay otro tipo de corona que puede verse rodeando focos luminosos lejanos, pero que nada tiene que ver con la difracción de la que estoy tratando. Esa tenue "corona endo-óptica" se crea en



Generación de una figura de difracción por el polvo en contacto con un espejo



Separación de los colores en una figura de difracción

el ojo del observador.) En cierta ocasión presencié un caso notable de anillos de Fraunhofer cuando me encontraba a un metro de distancia de un espejo empañado que recibía una iluminación intensa, procedente de una ventana situada a unos cinco metros detrás de mí: superpuestas a cada uno de los ojos de mi imagen en el espejo se veía una figura de Fraunhofer.

En los espejos cubiertos de polvo, el hecho de que la corona de difracción sea visible se debe a que la figura difractada por las partículas de polvo se refleja en la superficie posterior del espejo. Entonces, cualquiera que sea el ángulo de vista del observador, la figura de anillos y colores rodea siempre la imagen reflejada del manantial luminoso. Esta inamovilidad constituye la clave fundamental para distinguir los anillos de Fraunhofer de los anillos de Whewell-Quételet, que rodean la imagen únicamente si el observador y el manantial luminoso se encuentran en la misma perpendicular al espejo.

La nitidez de los anillos de Fraunhofer guarda una estrecha correlación con el tamaño de las partículas. Cuanto menores son éstas, mayor es la anchura angular de la figura de difracción. Por otro lado, una gran diversidad en el tamaño de las partículas produce una figura imprecisa porque se superponen las diferentes difracciones.

Para conseguir anillos de Fraunhofer bien definidos me valgo de una hoja de vidrio limpia o de un espejo ligeramente recubierto de polvos de licopodio. Las partículas de éste son de tamaño bastante uniforme y, en consecuencia, las figuras de difracción que generan son aproximadamente todas de las mismas dimensiones. Cuando ilumino una hoja de vidrio o un espejo cubiertos con la condensación de mi aliento, o de una ducha caliente, se forman unos anillos de Fraunhofer menos definidos (a veces, apenas se ven). Evidentemente, las gotitas que intervienen en la interferencia son de tamaños más diversos, y puede que se repartan más tupidamente de lo que suelen hacerlo las partículas de polvo, por lo que puede haber efectos interferenciales añadidos que oculten la sencilla figura de difracción que les he descrito.

La fotografía de las figuras de difracción que se ve en la ilustración superior de la página 127 se obtuvo con la luz de un láser de helio-neon, cuyo haz se hizo reflejar en un espejo cubierto de polvos de licopodio. La reflexión primaria iba a parar a una pantalla situada detrás del láser. Rodeando la superficie más brillante del conjunto se ve la figura de Fraunhofer; en la parte inferior de la fotografía, casi a la sombra del láser, está el centro de los anillos de Whewell-Quételet.

¿Se reflejaba realmente la figura, que yo he llamado figura de Fraunhofer, en la superficie posterior de un espejo? Para responderme a esta pregunta raspé parte del recubrimiento reflector del reverso de un espejo. Luego, empolvé ligeramente su superficie frontal con polvos de licopodio y lo iluminé. Los anillos de Fraunhofer aparecieron tan intensos como de costumbre en todas partes, salvo encima de la superficie no reflectora, donde se mostraban bastante débiles. Allí aún se veía una pálida imagen de la figura porque el reverso de la superficie no reflectora seguía reflejando alrededor del dos por ciento de la luz.

Entonces, para eliminar toda reflexión en la superficie raspada, rocié la superficie trasera del espejo con pintura negra mate. Una vez que ésta se secó, volví a empolvar la superficie frontal del espejo; en la superficie raspada no apareció la figura.

He meditado acerca de la razón por la cual haya de ser pequeño, o estar alejado del espejo, el manantial luminoso necesario para crear las dos figuras de interferencia que se ven en los espejos cubiertos de polvo. Cuando sostuve un manantial de luz intenso y extenso, como una bombilla eléctrica, cerca de un espejo cubierto de polvo, no pude observar figura alguna. La explicación de ello reside en dos características de la luz: la separación angular de los rayos y la coherencia de la luz lanzada sobre las partículas de polvo. Para que una pareja de rayos luminosos originen una figura de interferencia deben ser coherentes y aproximadamente paralelos en cada partícula. Si son incoherentes, la diferencia entre sus fases cambia de un instante a otro y también varía la intensidad de la interferencia cuando llegan al observador. Si no son paralelos, la figura de interferencia queda oculta por el solapamiento de las numerosas figuras que cada partícula emite bajo ángulos ligeramente distintos.

Un observador sagaz señalaría que las figuras de Fraunhofer creadas por las gotitas difieren en dos aspectos de las creadas por los polvos de licopodio. En el primer caso, el punto central aparece rodeado por una superficie oscura y, en el segundo, por una superficie clara. Por otra parte, las formas de las figuras difieren cuando la visual del observador no es perpendicular a la superficie del espejo o de la hoja de vidrio. Un observador atento observaría, además, que presentan una estructura granulosa intrigante, sobre todo cerca de la zona más luminosa. Estos y otros enigmas deben esperar otra oportunidad.

Libros

Autobiografía de Cajal, orígenes de la lengua catalana y el Tratado de astrología de Villena

Pedro Laín Entralgo, Antonio Tovar y Juan Vernet

RECUERDOS DE MI VIDA. HISTORIA DE MI LABOR CIENTÍFICA, por Santiago Ramón y Cajal. Madrid, Alianza Editorial, 1981. Los profesores A. Sols y F. Reinoso han tenido la excelente idea de publicar a través de Alianza Editorial, con la seguridad, por tanto, de que el libro va a llegar a muy extensos sectores de nuestros lectores cultos, la segunda parte de un libro tan famoso como poco conocido: los *Recuerdos de mi vida* de nuestro máximo hombre de ciencia. La parte primera de ese libro, relativa a la infancia y la juventud del autobiografiado, ha sido, sí, varias veces reeditada. No así esta parte segunda, consagrada a exponer sobria, precisa y documentadamente todo el curso de la investigación cajaliana, desde 1880, año en que se publica en Zaragoza un estudio experimental acerca de la génesis de la inflamación, hasta los trabajos de 1923, ya ulteriores a la jubilación académica del gran sabio. Con razón subraya Sols el penoso y ya remediado hecho de que sólo la fracción más anecdótica de *Recuerdos de mi vida* haya merecido, precisamente a partir de 1923, fecha de su tercera edición completa, la atención de los editores españoles. Veamos sumariamente las razones que tan acertada hacen a la reedición.

I. La importancia que la segunda parte de esos *Recuerdos* tiene hoy para el neurobiólogo ha sido enérgicamente afirmada en la "Introducción" de F. Reinoso: "Cajal sigue siendo uno de los autores neurobiológicos más citados... La importancia de *Histología del sistema nervioso del hombre y los vertebrados* es tal, como libro básico de consulta, que en uno de los laboratorios neurobiológicos americanos que en este momento disputa la primacía mundial, a todos los nuevos doctores en neurociencias se les obsequia con los dos tomos de la obra de Cajal... Si quisiéramos insistir sobre la influencia de la obra cajaliana en la ciencia neurológica actual, haríamos la historia completa de la neurobiología de los últimos noventa años; y es sólo una parte de la obra de Cajal la que se conoce bien." ¿Por qué todo esto? ¿Qué tiene esa ingente obra científica para que cien años después de iniciada —en 1881 publicó Cajal su primer trabajo neurohistológico: *Observaciones microscópicas sobre las terminaciones nerviosas en los músculos voluntarios de la rana*— siga gozando tan lozana actualidad? A mi modo de ver, en cinco puntos principales debe ser desglosada la respuesta.

1. La *fundamentalidad* de la construcción neurobiológica cajaliana. El centro de ella, la teoría neuronal del sistema nervioso, además de dar acabamiento a la teoría celular de Schleiden-Schwann-Virchow, como a lo largo de mi vida varias veces he dicho, afirmó un hecho y una interpretación que el progreso de las técnicas de observación y los avatares de la historia del pensamiento han convertido en base principal del saber neurológico. Como la obra conjunta de Mayer, Hehnholtz y Joule sigue vigente, porque atañe a un principio fundamental y todavía no controvertido de la ciencia de la naturaleza, así la de nuestro genial histólogo.

2. La insuperable *precisión objetiva* de las observaciones histológicas de Cajal. Las técnicas micrográficas ajenas, tal la de Golgi, mejoraban en sus manos, y las propias, así la del nitrato de plata reducido, daban en ellas el máximo rendimiento. Su mirada y su dibujo fueron, por otra parte, infaliblemente certeros. No puede así extrañar que sólo muy excepcionalmente hayan sido rectificados los "hechos" neuroanatómicos descubiertos por él.

3. La enorme *cantidad y variedad* de dichos descubrimientos. Poco después de la muerte del maestro, escribía el italiano E. Lugaro: "En el campo de la morfología nerviosa, puede decirse que Cajal, por sí solo, ha producido más que todos los otros neurólogos juntos." Así lo advertirá el lector del libro que ahora comento, y así lo hizo ver la documentada exposición de J. F. Tello en *Cajal y su labor histológica* (Madrid, 1935). "Cuando un aragonés se decide a tener paciencia, que le echen alemanes", dijo una vez, dando *sólo* por tenacidad baturra lo que en él era *también* genialidad. Juntas una y otra hicieron posible ese rotundo juicio de Lugaro.

4. La constante tendencia de su mente a la *interpretación funcional* de sus hallazgos, y el habitual acierto de sus interpretaciones. Adelantándose a tantos logros de la actual neurobiología, nadie como él y con tan escasos recursos materiales ha sabido reducir a unidad doctrinal la neuroanatomía y la neurofisiología. Polarización dinámica y axípeta del impulso nervioso, biología de las neurofibrillas, estructura funcional de la retina, revisión del problema anatomofisiológico del quiasma óptico, y en general de los entrecruzamientos en el sistema nervioso central, neurogénesis, degeneración y regeneración de los nervios... Una neurofisiología directamente basada en la neuroanatomía microscópica y precursora de lo que las técnicas microscópico-electrónicas, electrofisiológicas y neuroquímicas pronto habían de decir, se iniciaba con fuerza, ya desde 1888, en la obra histológica de Cajal.

5. El deliberado recurso de la investigación cajaliana al *método ontogénico y filogénico;* la metódica consideración de las estructuras neurológicas en etapas de su desarrollo embriológico en que su realidad es más sencilla que en el sistema nervioso del animal adulto —"Yo tuve la fortuna de contemplar por vez primera ese fantástico cabo del axón en crecimiento", dice una vez el sabio, con emoción contenida y muy legítimo orgullo— y la también metódica comparación de aquéllas en especies de distinto nivel en la evolución biológica: insectos, batracios, reptiles, aves, mamíferos.

Sumada la varia eficacia de estas cinco notas esenciales, se verá la razón del hecho bibliométrico citado por Reinoso: que Cajal sigue siendo uno de los autores neurobiológicos más citados. Y transparecerá, en consecuencia, la clave del vivo interés que para el neurofisiólogo actual poseen los *Recuerdos* ahora nuevamente editados.

II. Mas no sólo para el neurobiólogo es oportuna esta reedición; también lo es para el hombre de ciencia español, cualquiera que sea su dedicación, y hasta para el español a secas, si la cultura científica de su patria tiene para él algún valor. En dos sentidos alcanza esa oportunidad.

1. Ante todo, como demostración rotunda de que en España, si de ello hay resuelta voluntad individual y social, puede hacerse ciencia de calidad. Por Dios, no hablemos del "milagro de Cajal", aunque tan portentosa fuese la edificación de su obra en el semipáramo de nuestra vida científica, allá por los decenios centrales del siglo XIX. Esa obra fue, sí, genial y grandiosa, mas no un "milagro" otorgado por la especial providencia de los destinos españoles a una sociedad de vagos e insipientes. Sin mengua del evidente predominio de su carácter autodidáctico, sin detrimento de su vigorosa originalidad, la imponente construcción cajaliana recibió

incitaciones de su ambiente próximo, por modesto que éste fuera —baste citar los nombres de Maestre de San Juan y Simarro—, y pertenece históricamente al conjunto de la que cada uno en lo suyo levantaron otros hombres de ciencia españoles, algunos coetáneos del propio, Cajal, como Torres Quevedo, Olóriz, Gómez Ocaña, Turró, San Martín, Bolívar, Calderón, Quiroga y Rodríguez y Reyes Prósper, otros poco anteriores a él, entre ellos Macpherson, Torroja y García de Galdeano: los hombres que dedican la paz interior de la Restauración a la callada tarea de hacer, poca o mucha, verdadera ciencia al día. La conclusión es obvia: cuantas veces los hijos de Iberia se decidan a repetir por su cuenta lo que ese puñado de sus compatriotas hicieron —trabajar con ahínco, instalados en el nivel histórico de su tiempo—, habrá ciencia española internacionalmente estimable, aunque no toda ella alcance el cimero nivel a que llegó la histología de Cajal.

Dos estudios míos, "Cajal en la historia de España", en A. Albarracín Teulón y P. Laín Entralgo, *Santiago Ramon y Cajal o la pasión de España* (Editorial Labor, Barcelona, 1978), y "Cajal por los cuatro costados", en *Santiago Ramón y Cajal. Expedientes administrativos de grandes españoles* (Publicaciones del Ministerio de Educación y Ciencia, Madrid, 1978), exponen con relativa amplitud cuál fue el legado de Cajal a nuestra historia intelectual y social. A ellos remito al lector interesado por el tema. Ahora me limitaré a copiar unas líneas del primero: "¿Qué ha aportado Cajal a la historia de su pueblo? Cada cual dará la respuesta que considere oportuna. La mía dice así: una realidad, un mito y un problema." La realiad de su obra, de su ejemplo, de su escuela, de su personal modo de ser sabio español. Su condición mítica de varón glorioso por su ciencia y —en cierto modo— cómodo redentor de tantas deficiencias colectivas. "¡Es nuestro héroe!", decía con entusiasmo a Hugo Spatz, el gran histopatólogo alemán, un joven pensionado español. Y, por añadidura, la necesidad de replantear un viejo problema histórico, precisamente a la luz de lo que Cajal como científico hizo: después de él y de su repercusión directa e indirecta sobre la producción científica de los españoles, ¿cómo debe ser entendido y resuelto ese "problema de la ciencia española" que desde Juan de Cabriada hasta hoy mismo ha inquietado a todas nuestras conciencias sensibles?

2. Quisiera yo que la eficacia de estos *Recuerdos* cajalianos sobre el común de los españoles vocados a la ciencia, no sólo sobre los neurobiólogos, les llevase a una convicción inicial —"En España puede hacerse ciencia estimable, si de veras se quiere hacerla"— y una reflexión subsiguiente acerca del camino mejor para convertir en realidad esa convicción. Supuesta la existencia de talento suficiente, verdadera vocación y adecuada formación técnica en el aspirante a hombre de ciencia, pienso que, respecto del cumplimiento de tal empresa, hasta cuatro requisitos nos hace descubrir la autobiografía de Cajal. Muy concisamente voy a mencionarlos:

a. Ajustar el proyecto de trabajo a las posibilidades económicas de que se dispone. Mil y mil veces es preciso repetir que la atención presupuestaria de España a la producción de ciencia es literalmente deplorable; más aún, sonrojante. Muy cierto; pero si alguien quiere de veras investigar, pocas veces le faltará el mínimo suficiente.

b. Hallarse dispuesto a afrontar la escasez del interés social por la práctica y los resultados del trabajo científico. Transcribiré otra vez un expresivo texto de Marañón: "Cuando mi generación empezó a trabajar en sentido moderno, estábamos en la situación de Robinsón Crusoe, que tuvo que ser albañil, cazador, cocinero, maestro y público de sí mismo." Algo han cambiado a este respecto las cosas; pero, sin duda, no lo suficiente.

c. Ser capaz de elegir —quede aparte la capacidad innovadora del genio, nunca sujeta a reglas— un tema de investigación perteneciente a la ciencia que en aquel momento se hace. Por tanto: leer lo más actual entre lo que se publica y saber discernir en lo que se lee lo más prometedor.

d. Contar en primer término con las posibilidades de todo orden —no sólo, por tanto, las económicas— del país en que se vive. En nuestro caso, las de España.

Si sabe hacerse todo esto, y si se poseen en discreta medida los presupuestos antes señalados —vocación, talento, formación técnica—, es seguro que alguna ciencia se producirá, y así lo muestra la historia de nuestro país ulterior al que Cajal llama "mi año cumbre, mi año de fortuna". Incluso teniendo que soportar la penosa experiencia de ver cómo las lacras que el propio Cajal denunció en la organización de nuestra vida universitaria y científica todavía subsisten entre nosotros.

Con cierta emoción en el alma he releído el recuerdo que Cajal dedica a Victorino García de la Cruz, catedrático de química en la Universidad de Barcelona y compañero del gran sabio en la tertulia del Café de Pelayo. Victorino —"bonísima persona y talento clarísimo", se nos dice de él —fue en 1887 y 1888, cuando ante el microscopio y en la mente de Cajal estaba germinando la teoría de la neurona, el único confidente de los proyectos científicos cajalianos. "Al principio —leemos— me oía con extrañeza, casi con incredulidad. Patriota sincero, la desesperanza había ganado su espíritu y paralizado sus fuerzas. Mas al fin mis predicaciones obraron en él una especie de contagio, y siguiendo mi ejemplo, acabó por escoger en el dominio de la física, que cultivó siempre con amor, algunos temas de estudio *baratos,* es decir, accesibles a los medios con que contaba. Años después, recordando mis alentadoras exhortaciones, solía decir que sin mi estímulo no hubieran aparecido nunca sus interesantes descubrimientos sobre *Las leyes de los líquidos turbios y gases nebulosos,* y otras conquistas científicas de positivo valor." Trasladado a Madrid, Victorino García de la Cruz fue elegido miembro de la Real Academia de Ciencias —en su discurso de ingreso, *Leyes más sencillas y fundamentales que rigen la morfología interna de los materiales histológicos,* no es difícil percibir la influencia cajaliana— y murió en 1906, cuando por su edad, cincuenta y seis años, aún cabía esperar de él obra científica personal. ¿A cuántos españoles no podría ser aplicada como lección la pequeña historia de este que Cajal llama "pobre Victorino"?

Inevitablemente viene a mi memoria la dedicatoria de *Ideas sobre Pío Baroja,* de Ortega: "Hay seguramente unas cuantas docenas de jóvenes españoles que, hundidos en el oscuro fondo de la existencia provinciana, viven en perpetua y tácita irritación contra la atmósfera circundante... A esos muchachos díscolos e independientes, resueltos a no evaporarse en la ambiente impureza, dedico este ensayo..." Han pasado setenta años. No sólo en el oscuro fondo de la existencia provinciana, también en las aulas, las clínicas, los laboratorios y las bibliotecas de nuestras grandes ciudades hay docenas y docenas de jóvenes en los que, *mutatis mutandis,* perdurá el estado de ánimo entonces descrito por Ortega. Entre ellos quisiera yo que estuviesen los primeros lectores de este no envejecido, lozano, tonificante libro de Cajal. (P.L.E.)

LA FORMACIÓ DE LA LLENGUA CATALANA. ASSAIG D'INTERPRETACIÓ HISTÒRICA, por Antoni M. Badia i Margarit. Publicacions de l'Abadia de Montserrat, 1981. Un vol. de 207 páginas. El profesor Badía, conocido por sus estudios de filología románica y especialmente catalana, intenta en este volumen explicar la formación del catalán, particularmente en cuanto se refiere

a la diferenciación dialectológica en las dos variedades principales: oriental y occidental.

Sin desconocer la trascendencia de esta división para todo el dominio lingüístico de la lengua catalana, con la pertenencia de las variedades baleares al oriental y del territorio valenciano de lengua catalana al occidental, se concentra Badía en el Principado, y en él, como territorio originario, examina los factores geográficos que determinan en cierto modo la división, y revisa las teorías que han intentado explicar la formación de ambas zonas en hechos históricos y acaso étnicos.

Dos son, según Badía, tales teorías: la de Griera, que buscó la explicación en los centros vitales de la reconquista, y la de Sanchís Guarner, que supone en la población antigua y sus lenguas, es decir, en el sustrato, la causa de la diferenciación.

Griera acudió a una interpretación de la historia de la Marca hispánica como origen del dialecto oriental, y de Urgel como del occidental. La hipótesis de Griera tendría a su favor, como él mismo señaló, el paralelo con los otros centros de diferenciación dialectal a lo largo del norte de la Península: aragonés, castellano, leonés y asturiano, gallego y otros menores, correspondientes a entidades históricas de más o menos relieve. Sanchís Guarner buscó precisamente el punto débil en la teoría de Griera acudiendo al testamento de Wifredo el Velloso y buscando una explicación de la repartición del Condado entre sus tres o cuatro hijos de modo que en parte correspondería a los territorios dialectales: oriental y occidental, más las variedades pirenaicas (Badía, p. 19). Badía admite que el testamento del primer conde de Barcelona pudo tomar en cuenta la realidad lingüística, pero advierte que los hechos geográficos (determinantes en definitiva de la realidad lingüística) podrían suficientemente explicar la división.

Sin embargo, donde Badía encuentra la mayor objeción a la teoría de la reconquista como determinante es en el hecho de que el límite de la Cataluña nueva, al suroeste del Principado, cruza perpendicularmente el límite de los dos dialectos. Un hecho tan importante como la dirección de la reconquista de Ramón Berenguer I hacia el oeste, con el fin de debilitar a los moros de Lérida, no impidió que en esos territorios predominara, a pesar de la presencia política de Barcelona, el catalán occidental. Y del mismo modo en todos los territorios de reconquista al suroeste del Campo de Tarragona, como en todo el reino de Valencia, domina la variedad occidental, sin que lo impidan hechos como, por ej., la estadística del reparto de casas de Valencia, que se dieron en un 61 por ciento a catalanes orientales y un 38 a occidentales (sin contar los aragoneses). Badía ve, pues, que no es la reconquista lo determinante de la distribución dialectal.

Por ello se inclina en favor de Sanchís Guarner, que buscó otro factor para explicar la diferenciación: las etnias prerromanas. Ilergetes, ilercavones y edetanos, es decir, las regiones que actualmente corresponden al sur de las provincias de Tarragona y Lérida, reino de Valencia y sudeste de Aragón, son gentes iberas.

En cambio, para la Cataluña oriental Sanchís Guarner invoca las invasiones indoeuropeas señaladas por Bosch-Gimpera sobre bases arqueológicas como de campos de urnas célticos, y por Corominas sobre ciertas etimologías como sorotápticas o de túmulos. Admitiendo esto, frente al iberismo de la Cataluña nueva y Valencia, y frente al iberovasquismo de la cuenca del Segre y sus afluentes, tendríamos una Cataluña oriental indoeuropea.

Ahora bien, a estas explicaciones hay una razón cronológica decisiva que oponer. Las invasiones indoeuropeas que Bosch-Gimpera señaló en Cataluña son las más antiguas quizá de toda la Península, mientras que lo que sí podemos afirmar es que, en lo que sabemos por inscripciones y monedas, el nordeste de Cataluña, en los siglos inmediatamente anteriores a la conquista romana, y precisamente en tiempo de la romanización, pertenece al mundo ibérico, que se extiende, como es sabido, desde Andalucía hasta las avanzadas del sur de Francia en Narbona y Enserune. Las monedas de los indigetes de Ampurias se leen *Unticescen*, con la misma terminación que las de Urci (cerca de Almería) se leen *Urcescen*. Hace muchos años pude trazar un mapa de la difusión de la lengua ibérica por la distribución de unas cuantas características de este tipo, y el conjunto de monedas e inscripciones de Cataluña y del Rosellón Ibérico no se diferencia del resto (*Zephyrus* VII 81-83). Y unos años después, en 1964, frente a un intento de J. Maluquer de Motes y P. Pericay, "Problemas de la lengua indígena de Cataluña", en el volumen *Problemas de la prehistoria y de la arqueología catalanas, II Symposium de prehistoria peninsular,* Universidad de Barcelona 1963, 101-143, nos oponíamos a reconocer elementos indoeuropeos en el ibérico de Cataluña, que en cuanto se puede afirmar de una lengua no descifrada, pero bastante documentada, muestra palabras, desinencias y elementos de toda clase idénticos a los del resto del territorio (v. mi reseña en *Zephyrus* XV 131 ss.).

Los indoeuropeos en Cataluña a que se refiere Bosch-Gimpera son del año 900 a. C., y arqueológicamente el mismo autor los reconocía como subsistentes en ciertas regiones hasta el año 500. Pero los topónimos célticos más seguros en Cataluña (los en *-dunum* y en *-acum,* que invocaba el mismo autor) pertenecen sin duda al tipo galo, que no se desarrolla antes de la época de la Tène, es decir, a partir del siglo V a. C. Los invasores sorotápticos de Corominas no son un concepto arqueológico, sino una hipótesis para explicar ciertos elementos indoeuropeos en la onomástica peninsular (y no sólo en Cataluña) que podrían calificarse de "preceltas". Pero la escasez de esos testimonios no permite en modo alguno sostener que una hipotética población en épocas remotas, en la primera mitad del I milenio a. C., pueda determinar los rasgos del latín que predominó en Cataluña a partir del siglo II a. C. Los antiguos pobladores de la Cataluña oriental que emergen en las más antiguas noticias históricas: ausetanos, laietanos, indigetes, ceretanos, berguistanos y demás, no se diferencian en nada seguro de sus vecinos del este y del sur en cuanto nos enseñan las inscripciones. En ellas nada se encuentra de indoeuropeo, que se habría perdido, como más tarde se perdieron el godo o el franco en zonas de romanización muy fuerte.

En este punto, de que el latín de Cataluña oriental fuera recibido y adoptado por gentes de lengua indoeuropea, nos encontramos llenos de reservas ante la tesis que, basada principalmente en M. Tarradell, sostiene Badía. Para nosotros no hay ninguna prueba de que lenguas indoeuropeas se mantuvieran en Cataluña y Rosellón en las tardías épocas de la implantación del latín. No admitiríamos más excepción que los elementos célticos galos, que es cosa aceptada por los historiadores que no llegan a la costa francesa del Mediterráneo antes del siglo III (así tenemos en el Rosellón una moneda con la leyenda ibérica *Biricantin* que ha de corresponder a los *Brigantes* históricamente conocidos, pero tal fenómeno no se da en Cataluña al sur de los Pirineos).

Los estudios de Corominas sobre los restos lingüísticos prerromanos en Cataluña no permiten sostener la persistencia de antiguas lenguas indoeuropeas hasta época tardía. Reseñando el libro de U. Schmoll reconocía hace veinte años Corominas (*Tópica hespérica* II 247) que la línea divisoria que yo había establecido entre la Hispania indoeuropeizada y la no indoeuropea era segura. Y en su estudio sobre los plomos de Arles (Amélie-les-Bains), en el que busca hallar precisamente elementos de una lengua "sorotáptica", lo más seguro y convincente en los que no es latín es

la palabra vasca *Neskas,* que es la designación de las ninfas de la fuente medicinal. En poco se diferencia este mundo lingüístico rosellonés del que el propio Corominas ha estudiado al oeste en una inscripción, como él dice, "en basc ribagorçà" del siglo I d. C. Es decir, que en la Cataluña oriental, como en la occidental, nos hallamos en el mismo ambiente lingüístico ibérico de toda la España oriental, con la presencia de elementos vascos en el Pirineo (me remito a la presentación que Corominas ha hecho en catalán de estos dos estudios suyos recientes en la obra *Entre dos llenguatges,* II, 132 ss. y 142 ss.).

Evidentemente, si no creemos en fuerzas telúricas o en fantasías raciales, un sustrato no obra más que a través de una situación bilingüe durante una o varias generaciones. Los indoeuropeos que llegaron a Cataluña, y a otras regiones peninsulares, en el siglo VIII, supongamos, y de los que no hay testimonios lingüísticos, no pudieron influir en la adaptación local del latín. Los romanos encontraron en Cataluña pueblos ibéricos de lengua, y seguramente vascos en las regiones montañosas del oeste. La diferenciación del catalán oriental y occidental puede basarse en la presencia de esos "vascos ribagorzanos" que Corominas, siguiendo a R. d'Abadal, encuentra al este, pero no en unos indoeuropeos que lingüísticamente no han dejado ningún testimonio. Un latín profundamente arraigado, precisamente como Badía muestra en su estudio, basta seguramente para explicar la formación de una lengua románica muy semejante por cierto a la del también profundamente latinizado mediodía de Francia.

Más fácil nos es seguir al profesor Badía en sus explicaciones históricas: cuando examina los matices de la romanización, y contrapone, sobre pruebas seguras, la situación de la Cataluña oriental con la de Pallars y Ribagorza. En la toponimia y la arqueología busca acertadamente la demostración de la densidad de la romanización, que luego se continúa en la época visigoda con los nombres de santos en la toponimia.

Defiende Badía justificadamente la unidad de los dos dialectos catalanes principales. Los matices del catalán oriental, con su vocalismo más alejado del castellano (donde pesa un sustrato vasco, como en el catalán occidental), se pueden explicar con el mayor arraigo de la colonización latina. Una afinidad indoeuropea para explicar los rasgos del catalán oriental nos parece inadmisible, pues se trata de una hipótesis muy débil ante el hecho innegable del iberismo de cuanto sabemos de la lengua antigua de Cataluña. El mismo Badía expone (p. 110) sus dudas sobre la teoría del sustrato, que evidentemente no es admisible sino cuando generaciones bilingües forman la mezcla en que sobrevive la lengua desaparecida.

Una valiosa contribución original de Badía es la de considerar los árabes como elemento destructor de la romanización. La presencia de onomástica árabe es por sí prueba de desromanización. Y con estudios onomásticos se ve que la zona occidental de la Cataluña nueva ha sido más profundamente desromanizada que la oriental. La comparación que Badía hace del elemento germánico en Galia como desromanizador ilustra la significación de los arabismos en la Cataluña nueva.

Novedades importantes son las que contiene el breve capítulo sobre el balear, donde la instalación de un catalán oriental con peculiaridades locales es un problema con varias incógnitas. En conjunto la nueva obra de Badía, que constituye la más completa exposición sobre los orígenes del catalán, aporta teorías que son más aceptables cuando se fundan en realidades históricas más próximas que cuando invocan una hipótesis más dificilmente comprobable. (A.T.)

TRATADO DE ASTROLOGÍA ATRIBUIDO A ENRIQUE DE VILLENA. Edición y notas, Pedro M. Cátedra. Introducción, Julio Samsó. Edición patrocinada por Río Tinto Minera S.A., Barcelona, 1980. La idea de patrocinar ediciones de clásicos de la historia de la ciencia por grandes sociedades industriales españolas no es nueva, pero hay que confesar que no ha cundido mucho. Conozco una media docena de libros que se han impreso bella y cuidadosamente por estas entidades con el fin de regalarlos durante las fiestas navideñas a sus accionistas. El *Tratado de Astrología,* del que ahora vamos a hablar, tiene este mismo origen: es una hermosa edición de bibliófilo del manuscrito único conservado en la Biblioteca Nacional de Madrid, que si bien había sido utilizado esporádicamente por los investigadores (v. g. J. M. Millás Vallicrosa, a cuya memoria se dedica esta edición), nunca había visto la luz pública como un todo. Y es una lástima puesto que el libro en cuestión demuestra que el autor del mismo —fuera o no Enrique de Villena, tema que se discute en la Introducción— era un español muy culto, tal vez no excesivamente ducho en el campo de las matemáticas y de la astronomía pero que conocía muy bien cuáles eran las corrientes dominantes en el mundo científico de la época y sabía divulgarlas con bastante claridad.

La obra consta de dos partes claramente diferenciadas: *1)* La transcripción del texto debida a P. M. Cátedra, establecida con todo el rigor que exigen las modernas normas paleográficas para la edición de textos medievales y *2)* un exhaustivo estudio científico del profesor Samsó, miembro de la Académie Internationale d'Histoire des Sciences (París) y al cual se deben muchos trabajos sobre la evolución de la ciencia medieval. La primera parte, la Introducción, comprende las páginas 13-68 y a ella voy a referirme substancialmente; la segunda — edición del texto y notas al mismo de tipo paleográfico o de identificación de fuentes por lo general literarias— va de la página 71 a la 181 y lleva intercaladas bastantes láminas a todo color —las que figuran en el manuscrito— que pueden entenderse viendo la interpretación moderna que de las mismas hace Samsó en su Introducción.

Samsó ha analizado el texto con todo cuidado y ha reunido bajo epígrafes muy concretos todos los materiales que, sobre un mismo tema, se encuentran esparcidos a lo largo del manuscrito haciendo los reenvíos necesarios a los folios del mismo. Así consigue dar al lector moderno, en lenguaje moderno, una visión clara del contenido del *Tratado* y en algunos puntos en que éste se muestra confuso ha procedido a avanzar sus propias hipótesis de interpretación avaladas por el control del cálculo. Igualmente, y siempre que ha podido, ha señalado las fuentes científicas utilizadas por el Autor y que van desde la doctrinas helenísticas del *Almagesto* y *Tetrabiblos,* de Tolomeo, a textos indios, latinos (San Isidoro) y algunos procedentes de autores indeterminados, pero que no por eso dejan de tener un interés máximo —lo veremos al tratar del tamaño del Universo— para comprender el ambiente en que se movió, a fines del siglo XV, es decir, unos setenta y cinco años después de escrito el *Tratado,* el propio Copérnico.

Un científico moderno que lea el texto prescindiendo de la Introducción encontrará una serie de pasajes comprensibles con sus conocimientos actuales. Por ejemplo, cuando se trata del calendario eclesiástico o se nos expone la teoría de los terremotos o los valores numéricos de distintos ciclos planetarios. En lo que sigue, vamos a prescindir de analizar estos pasajes, y, en cambio, vamos a comentar algunos detalles del *Tratado* algo más crípticos siguiendo el orden en que aparecen en la Introducción.

Tras una exposición del estado de la astronomía, astrología y ciencias conexas a lo largo de la Edad Media (págs. 14-21) en la que señala que, según el testimonio de Llull, el único instrumento utilizado a bordo de los barcos en aquel entonces era el nocturlabio (aparato para determinar la hora de la

noche en alta mar (pág. 16 *a*), apunta unas páginas más adelante (20 *a*) siguiendo a Poulle que, cuando menos teóricamente, se conocía desde la época de Raymond de Marsella (fl. c. 1140) el manejo a bordo del astrolabio para determinar la latitud. Problema éste que, junto al de la determinación de los acimutes, debió preocupar a nuestros marinos medievales, quienes debían establecer el rumbo mediante la brújula o mediante tablas de declinación del Sol y alturas de éste del tipo de las calculadas por Ibn Yunus (m. 1009, Cf. King (1973)) y que nos parece evidente que eran conocidas en la España del siglo XV (Cf. pág. 21 *a*), aunque tal vez —lo dudo— no supieran utilizarse correctamente en la mar.

La justificación del título de la obra y la identificación de sus fuentes cubre las páginas 21-26, siguiendo a continuación el epígrafe titulado "La estructura y las dimensiones del Universo" (págs. 26-32) que, para mí, justifica uno de los factores ideológicos causantes de la reforma copernicana. Prescindamos de la estructura y de los modelos cinéticos del movimiento de los planetas (págs. 53-67) para fijarnos exclusivamente en las ideas que, en el siglo XV, se tenían sobre el tamaño del Universo. Samsó subraya, correctamente, que el universo del *Tratado*, cualesquiera que sean las discrepancias numéricas que presenta en los distintos pasajes en que alude a su radio, es enormemente mayor que el fijado por Tolomeo y que estas discrepancias remontan a un texto atribuido a Avicena (m. 1037) y a otro de Leví ben Gerson (m. 1344). El cómo llegó este último a fijar que las estrellas fijas se encontraban en una distancia de $159 \times \times 10^{12} + 6515 \times 10^{8} + 1338 \times 10^{4} + 944$ radios terrestres (R) nos escapa. Pero lo que no se nos escapa es que, desde la más remota antigüedad, existió una tendencia constante a ampliar los límites del Cosmos y que autores clásicos, nada menos que Aristarco, fíjanse [Arquímedes, *Arenario apud Bausani,* (1977), 206-207] la distancia de la Tierra al Sol en 185×10^{11} kilómetros y el del Universo a 9250×10^{12}, es decir, a un año luz, mucho antes de que Tolomeo lo redujera a 20160 R. Después de la muerte de Tolomeo siguieron coexistiendo distintos valores para el tamaño del Cosmos. Las *Epístolas* árabes de los Hermanos de la Pureza (c. 1000) colocan (Bausani, 1977 y 1978 pp. 103-110 nota) las estrellas fijas a 67878 R, mientras que su coetáneo —y mucho más entendido que ellos en cuestiones astronómicas, al-Battaní (m. 929), cuyas medidas de paralajes pudieron haberle llevado a otras conclusiones (Swerdlow, 1972), no se atrevió a discrepar públicamente de Tolomeo. Unos siglos después, el persa Qutb al-Din al-Sirazí (m. 1311) fijó la distancia de las estrellas fijas a la Tierra en 140147 R y nuestro *Tratado* —y conviene repetir que refleja las ideas de un hombre culto de su época, aunque no de un especialista— en una de sus apreciaciones de los confines del Universo la establece, según Samsó, en 5256×10^{8} millas. Lo cual, tomando como valor aproximado de la milla árabe el de 1,8 kilómetros, nos lleva a un radio enorme de la esfera celeste.

Pero hay más: Costabel, en el simposio del centenario de Copérnico de Toruń (1973), señaló que en la Colegiata de Champeaux, cerca de Melun, en una de las ménsulas del coro, aparecía (1525) Dios creador cogiendo la bola del mundo y echándola a rodar, es decir, dotando a nuestro planeta de movimiento de rotación y traslación. Es cierto que el escultor de la misma pudo estar ya influido por el *Commentariolus* de Copérnico, aunque creo que nunca lo sabremos. Pero de lo que no cabe duda es de que éste conoció valores del radio del Universo mucho mayores que el admitido por Tolomeo y, oficialmente, por él mismo (en sus años de estudiante en Italia aparte de conocerse varios de los valores antecitados debió enterarse también de los dados por el *Arenario* y puestos a nombre de Aristarco), razón por la cual podía construir, sin confesar los motivos, su nuevo sistema celeste haciendo omisión de las críticas que pudieran hacérsele por la falta aparente de paralaje de las estrellas fijas.

Samsó analiza a continuación los conocimientos meteorológicos y geográficos expuestos en el *Tratado* de modo muy somero —no merecen más— y a continuación se adentra en la exposición de los datos astrológicos (págs. 35-48) que encuadra bien dentro de la tradición imperante en la época exponiendo, de paso, anécdotas y doctrinas poco o nada conocidas hasta ahora pero que subyacen en la obra, la cual, sin su previo estudio, sería ininteligible para el lector corriente. Igualmente detalla los conocimientos trigonométricos, bastante elementales, del autor analizando sus distintas procedencias y cierra su Introducción con el estudio de las teorías planetarias en boga en la primera mitad del siglo XV.

Así, una obra aparentemente anodina —que no lo es— se transforma en una fuente de información de primer orden para saber el estado de las ciencias exactas en los orígenes del Renacimiento y aclara una serie de puntos que, hasta hoy, aparecían confusos.

Es de esperar que Samsó y Cátedra reediten este libro en una edición popular que lo haga accesible a un más amplio número de lectores. (J. V.)

Bibliografía

Los lectores interesados en una mayor profundización de los temas expuestos pueden consultar los trabajos siguientes:

ARMAS CON GUIA DE PRECISION

NEW TECHNOLOGY AND MILITARY POWER: GENERAL PURPOSE MILITARY FORCES FOR THE 1980S AND BEYOND. Seymour J. Deitchman. Westview Pres, 1979.

THE FY1981 DEPARTMENT OF DEFENSE PROGRAM FOR RESEARCH, DEVELOPMENT, AND ACQUISITION. Subsecretario de defensa para la investigación e ingeniería William J. Perry. Government Printing Office, 1980.

JANE'S WEAPON SYSTEMS 1980-81. Dirigido por Ronald T. Pretty. Franklin Watts Inc., 1980.

THE IMPACT OF NEW MILITARY TECHNOLOGY. Dirigido por Jonathan Alford. Gower Publishing, 1981.

LAS ESTRELLAS MAS RECIENTES DE ORION

INTERSTELLAR MOLECULES. Dirigido por B. H. Andrew. D. Reidel Publishing Company, 1980.

IINFRARED ASTRONOMY. Dirigido por C. G. Wynn-Williams y D. P. Cruikshank. D. Reidel Publishing Company, 1981.

OUTFLOW OF MATTER IN THE KL NEBULA: THE ROLE OF IRc2. D. Downes, R. Genzel, E. E. Becklin y C. G. Wynn-Williams en *The Astrophysical Journal*, vol. 244, n.º 3, 1.ª parte, págs. 869-883; 15 de marzo de 1981.

LA ENFERMEDAD HOLANDESA DEL OLMO

ATTRACTANT PHEROMONE OF THE EUROPEAN ELM BARK BEETLE (*SCOLYTUS MULTISTRIATUS*): ISOLATION, IDENTIFICATION, SYNTHESIS AND UTILIZATION STUDIES. G. N. Lanier, R. M. Silverstein y J. W. Peacock en *Perspectives in Forest Entomology*, dirigido por John F. Anderson y Harry K. Kaya. Academic Press, 1976.

SOME PHYTOTOXIC GLYCOPEPTIDES FROM *CERATOCYSTIS ULMI*, THE DUTCH ELM DISEASE PATHOGEN. Gary Strobel, Neal Van Alfen, Kenneth D. Hapner, Michael McNeil y Peter Albersheim en *Biochimica et Biophysica Acta*, vol. 538, n.º 1, págs. 60-75; 3 de enero de 1978.

PHEROMONE-BAITED TRAPS AND TRAP TREES IN THE INTEGRATED MANAGEMENT OF BARK BEETLES IN URBAN AREAS. G. N. Lanier en *Management of Insect Pests with Semiochemicals: Concepts and Practices*, dirigido por E. E. Mitchell. Plenum Press, en prensa.

EL RIBOSOMA

ESTUDIOS SOBRE LOS RIBOSOMAS MEDIANTE DISPERSIÓN DE NEUTRONES; Donald M. Engelman y Peter B. Moore en *Investigación y Ciencia*, diciembre, 1976.

RIBOSOME STRUCTURE AND tRNA BINDING SITES. James A. Lake en *Transfer RNA: Structure, Properties and Recognition*, dirigido por Paul R. Schimmel, Dieter Söll y John N. Abelson. Cold Spring Harbor Laboratory, 1979.

LA FORJA DE LOS METALES

FORGING MATERIALS AND PRACTICES. A. M. Sabroff, F. W. Boulger y H. J. Henning. Reinhold Book Corporation, 1968.

FORGING EQUIPMENT, MATERIALS, AND PRACTICES. Dirigido por Taylan Altan. Subvencionado por el Air Force Materials Laboratory. The Metals and Ceramics Information Center (Handbook No. MCIC-HB-03), Battelle Memorial Institute, Columbus Laboratories, Metalworking Division, 1973.

THE DEFORMATION OF METALS AT HIGH TEMPERATURES. Hugh J. McQueen y W. J. McGregor Tegart en *Scientific American*, vol. 232, n.º 4, págs. 116-125; abril, 1975.

LOS MONZONES

MONSOON METEOROLOGY. Número especial de *Pure and Applied Geophysics*, dirigido por T. N. Krishnamurthi, vol. 115, n.os 5 y 6, págs. 1087-1529; 1977.

THE SCIENCE AND WONDERS OF THE ATMOSPHERE. Stanley D. Gedzelman. John Wiley & Sons, Inc., 1980.

MONSOON DYNAMICS. Dirigido por Sir James Lighthill y R. Pearce. Cambridge University Press, 1981.

EL PROGRAMA DEL ALCOHOL EN BRASIL

ETANOL: COMBUSTÍVEL E MATÉRIA PRIMA. Stumpf, U. E. Semana da Tecnologia Industrial. Ministerio da Industria e Comércio, Secretaria da Tecnologia Industrial. Rio de Janeiro, 1976.

ALCOHOL FROM PLANT PRODUCTS. A BRAZILIAN ALTERNATIVE TO THE ENERGY SHORTAGE. Goldemberg, J. y Moreira, J. R. Preimpresión IFUSP/P-126, Instituto de Física da Universidade de São Paulo. São Paulo, Brasil.

ENERGY BALANCE FOR THE PRODUCTION OF ETHYL AND METHYL ALCOHOL. Moreira, J. R., Vanin, V. R. y Goldemberg, J. Workshop on Fermentation Alcohol for Use as Fuel and Chemical Feedstock in Developing Countries, Viena, 26-30 de marzo de 1979. Preimpresión IFUSP/P-165, Instituto de Física da Universidade de São Paulo; 1979.

SADI CARNOT

A SHORT HISTORY OF THE STEAM ENGINE. Henry W. Dickinson. Cambridge University Press, 1963.

FROM WATT TO CLAUSIUS: THE RISE OF THERMODYNAMICS IN THE EARLY INDUSTRIAL AGE. D. S. L. Cardwell. Cornell University Press, 1971.

REFLECTIONS ON THE MOTIVE POWER OF FIRE. Sadi Carnot. Peter Smith, 1977.

RÉFLEXIONS SUR LA PUISSANCE MOTRICE DU FEU. Sadi Carnot. Librairie Philosophique J. Vrin, 1978.

JUEGOS MATEMATICOS

GALILEO'S DISCOVERY OF THE PARABOLIC TRAJECTORY. Stillman Drake y James MacLachlan en *Scientific American*, vol. 232, n.º 3, págs. 102-110; 1975.

DO SIMILAR FIGURES ALWAYS HAVE THE SAME SHAPE? Paul G. Kumpel, Jr., en *Mathematics Teacher*, vol. 68, n.º 8, págs. 626-628; diciembre, 1975.

TALLER Y LABORATORIO

INTERFERENCE IN SCATTERED LIGHT. A. J. de Witte en *American Journal of Physics*, vol. 35, n.º 4, págs. 301-313; abril, 1967.

THE CORONA. R. A. R. Tricker en *Introduction to Meteorological Optics*. American Elsevier Publishing Company, Inc., 1970.

THE CORONA. Robert Greenler en *Rainbows, Halos, and Glories*. Cambridge University Press, 1980.

¿Nos hallamos en la frontera de una nueva revolución industrial? De la mano de la física vinieron las dos primeras, representadas por la máquina de vapor y los ordenadores. ¿Será biológica la tercera?

MICROBIOLOGIA INDUSTRIAL E INGENIERIA GENETICA

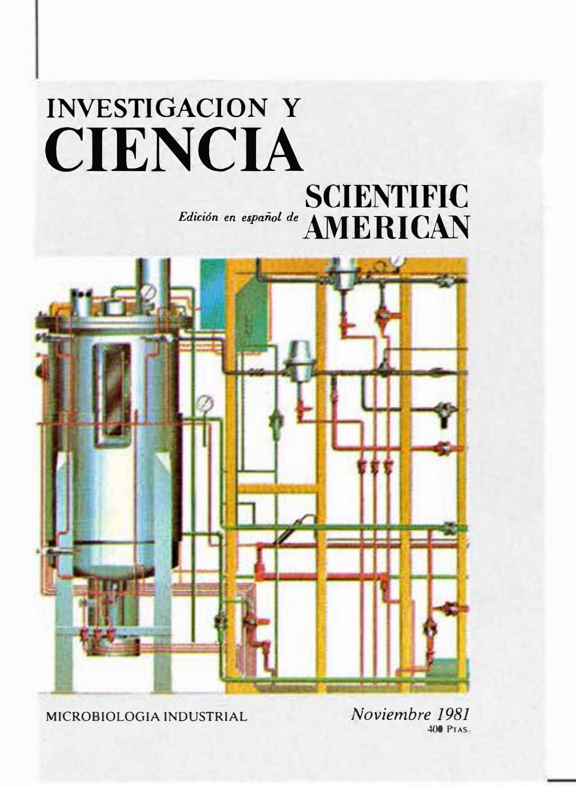
Este será el tema monográfico del próximo número extraordinario de INVESTIGACION Y CIENCIA, correspondiente a noviembre de 1981.

A lo largo de los últimos años se ha venido asistiendo a unos avances espectaculares en la identificación de las unidades fundamentales de la vida y de los procesos que atraviesan. Del conocimiento se ha pasado ya a la manipulación y al aprovechamiento industrial de sus posibilidades.

Quienes se encuentran en la avanzadilla de ese nuevo mundo de la tecnología e ingeniería genética explican y muestran, en un texto e ilustraciones únicos, cuál es la clave que promete cambiar el futuro económico de no pocos países y sectores industriales.

Un poder no exento de riesgos ni de reparos éticos porque no sólo afecta a los científicos y a los políticos sino a toda la sociedad.

Durante muchos años este volumen único será punto obligado de referencia por recogerse en él cuanto de más autorizado puede encontrar el lector en ese nuevo mundo del aprovechamiento de la biología en la industria y el desarrollo.



ARTICULOS

- **Microbiología industrial**
- **¿Qué es la ingeniería genética?**
- **Microorganismos industriales**
- **La programación genética de microorganismos industriales**
- **La producción microbiológica de alimentos y bebidas**
- **La producción microbiológica de productos farmacéuticos**
- **La producción microbiológica de productos químicos industriales**
- **Métodos de producción en microbiología industrial**
- **Biología molecular en agricultura**
- **La inhibición de la síntesis de proteínas**